NCS TIB 93-2

NATIONAL COMMUNICATIONS SYSTEM

TECHNICAL INFORMATION BULLETIN 93-2

IMPLICATIONS OF HDTV FOR GOVERNMENT TELECOMMUNICATIONS

JANUARY 1993

OFFICE OF THE MANAGER
NATIONAL COMMUNICATIONS SYSTEM
701 SOUTH COURT HOUSE ROAD
ARLINGTON, VA 22204-2198

THIS PAGE WAS INADVERTENTLY LEFT OUT OF THE ORIGINAL REPORT

PER JANET ORNDORFF
NATIONAL COMMUNICATIONS SYSTEM
(703) 607-6204

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 1993	Final Report

4. TITLE AND SUBTITLE

Implications of HDTV for Government Telecommunications

5. FUNDING NUMBERS

DCA100-91-C-0031

6. AUTHOR(S)

Stephen Perschau

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Delta Information Systems, Inc.
300 Welsh Road, Bldg 3, Suite 120
Horsham, PA 19044-2273

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Communications System
Office of Technology and Standards Division
701 South Court House Road
Arlington, Virginia 22204-2198

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NCS TIB #93-2

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In recent years, there has been considerable activity in the development of technology and standards related to High Definition Television (HDTV). The broad purpose of this task is to analyze the HDTV standardization process from the perspective of the telecommunications interests of the Federal Government. To accomplish this objective, the work on the project was divided into the four tasks. Analyze activities which are contributing directly to HDTV standards, analyze activities which are contributing to telecommunication standards which are related to HDTV, analyze potential applications of HDTV in the Federal Government, and identify and recommend actions which (1) maximize the interoperability between HDTV systems and government telecommunications, and (2) maximize the applicability of HDTV within the government community. This document summarizes the work performed on each of these four tasks.

14. SUBJECT TERMS

Improved Definition Television (IDTV), Extended Definition Television (EDTV), High Definition Television (HDTV), Advanced Digitial Television (ADTV)

15. NUMBER OF PAGES

300

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASS	UNCLASS	UNCLASS	UNLIMITED

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet *optical scanning requirements*.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

- C - Contract
- G - Grant
- PE - Program Element
- PR - Project
- TA - Task
- WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- DOE - See authorities.
- NASA - See Handbook NHB 2200.2.
- NTIS - Leave blank.

Block 12b. Distribution Code.

- DOD - Leave blank.
- DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- NASA - Leave blank.
- NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

NCS TECHNICAL INFORMATION BULLETIN 93-2

IMPLICATIONS OF HDTV FOR GOVERNMENT TELECOMMUNICATIONS

PROJECT OFFICER

GARY REKSTAD
Electronics Engineer
Office of Technology
and Standards

APPROVED FOR PUBLICATION:

Dennis Bodson

DENNIS BODSON
Assistant Manager
Office of Technology
and Standards

FOREWORD

Among the responsibilities assigned to the Office of the Manager, National Communications System, is the management of the Federal Telecommunication Standards Program. Under this program, the NCS, with the assistance of the Federal Telecommunication Standards Committee identifies, develops, and coordinates proposed Federal Standards which either contribute to the interoperability of functionally similar Federal telecommunication systems or to the achievement of a compatible and efficient interface between computer and telecommunication systems. In developing and coordinating these standards, a considerable amount of effort is expended in initiating and pursuing joint standards development efforts with appropriate technical committees of the International Organization for Standardization, and the International Telegraph and Telephone Consultative Committee of the International Telecommunication Union. This Technical Information Bulletin presents an overview of an effort which is contributing to the development of compatible Federal, national, and international standards in the area of High Definition Television as related to telecommunications. It has been prepared to inform interested Federal activities of the progress of these efforts. Any comments, inputs or statements of requirements which could assist in the advancement of this work are welcome and should be addressed to:

Office of the Manager
National Communications System
Attn: NT
701 S. Court House Road
Arlington, VA 22204-2198

IMPLICATIONS OF HDTV
FOR GOVERNMENT TELECOMMUNICATIONS

December, 1992

FINAL REPORT
Task No. 6
DCA100-91-C-0031

Submitted to:
NATIONAL COMMUNICATIONS SYSTEM
WASHINGTON, DC

D E L T A I N F O R M A T I O N S Y S T E M S, I N C.
300 Welsh Road, Ste. 120
Horsham, PA 19044-2273
TEL: (215) 657-5270 FAX: (215) 657-5273

TABLE OF CONTENTS

APPENDIX B - ANALYSIS OF INTEROPERABILITY EVALUATIONS BY PROPONENT AND REVIEW BOARD

APPENDIX C - PS-WP/4 FINAL REPORT

APPENDIX D - SMPTE HEADER/DESCRIPTOR TASK FORCE: FINAL REPORT

APPENDIX E - REPORT OF THE TASK FORCE ON DIGITAL IMAGE ARCHITECTURE

APPENDIX F - STATEMENT CONCERNING USER REQUIREMENTS FOR SOURCE CODING AND MULTIPLEXING FOR BROADCAST APPLICATIONS

1.0 OVERVIEW

This document summarizes work performed by Delta Information Systems, Inc. (Delta) for the National Communications System (NCS), Office of Technology and Standards. The NCS is responsible for the management of the Federal Telecommunications Standards Program, which develops telecommunications standards, whose use is mandatory for all Federal departments and agencies.

This document is a final report for a Task Order on Contract DCA100-91-C-0031. The titles for the contract and Task Order are listed below.

- Contract DCA100-91-C-0031
  Development of Federal Telecommunication Standards Relating to Digital Facsimile and Video Teleconferencing
- Task 2/Subtask No. 6 (1992)
  High Definition Television

In recent years, there has been considerable activity in the development of technology and standards related to High Definition Television (HDTV). The broad purpose of this task is to analyze the HDTV standardization process from the perspective of the telecommunications interests of the Federal Government. To accomplish this objective, the work on the project was divided into the four tasks listed below.

1. Analyze activities which are contributing directly to HDTV standards,
2. Analyze activities which are contributing to telecommunication standards which are related to HDTV,
3. Analyze potential applications of HDTV in the Federal Government,
4. Identify and recommend actions which (1) maximize the interoperability between HDTV systems and government telecommunications, and (2) maximize the applicability of HDTV within the government community.

This document summarizes the work performed on each of these four tasks and an overview on each task is provided below.

Status of HDTV Standards

- The FCC and ACATS organizations are evaluating five proposed HDTV systems, four of which are all-digital. It is a virtual certainty that one of the four all-digital systems (or a hybrid combination) will be selected to be the winner in 1993.

- On the international front, the Japanese are providing the MUSE system (1125 lines; 60 fields/sec.) on an operational tariffed basis, while the Europeans are experimenting with the MAC system (1250 lines; 50 fields/sec.). While these systems are both fundamentally analog in nature, there is considerable concern in these foreign countries that the U.S. all-digital standard will supersede their systems. For this reason, they are supporting work by the CCIR to develop an international all-digital HDTV standard.

Status of Telecommunication Standards Related to HDTV

- The ISO organization is developing the international MPEG-2 standard which is a toolkit having several important application profiles -- (1) HDTV, (2) H.26x (video codec for ATM/BISDN), (3) basic NTSC entertainment TV (contribution, distribution, VCR), and (4) cable/satellite systems. The specification for the MPEG-2 standard is scheduled to be frozen in March, 1993. Significant effort is being made to align the FCC HDTV standard with MPEG-2 to maximize interoperability.

- The CCITT ATM Video Coding Experts Group is developing Recommendation H.26x to be compatible with MPEG-2. This will insure the ability of an extension of H.26x to efficiently transmit the HDTV signal over the future ATM/BISDN network.

HDTV Applications in the Federal Government

It is generally acknowledged that the use of television throughout the

Federal government is pervasive and important to its successful operation. When the domestic TV standard was upgraded from monochrome to color, there was a major breakthrough for the Federal government. It is anticipated that the conversion from NTSC to HDTV will be at least as revolutionary. The higher resolution will provide a far more effective link between Federal employees and the vast quantities of information which is generated throughout the Federal government on an ever-expanding scale. Key application areas include multimedia desktop workstations, training/distance learning, simulation, command/control, reconnaissance/surveillance, and video teleconferencing.

Recommendations

Based upon the work performed on this project, the following recommendations are proposed.

- To maximize interoperability between future HDTV systems and Federal telecommunication systems (e.g., video teleconferencing), it is recommended that the domestic HDTV standardization process and the international MPEG-2 standardization efforts be aligned as closely as possible. To accomplish this alignment, it is recommended that Delta continue its active participation in both of these organizations. To insure that the H.26x Broadband ISDN video coding standard is closely aligned with these standards (HDTV, MPEG-2), it is further recommended that Delta continue its activity with the ATM Video Coding Experts Group. For example, it is important that the MPEG-2 standard have the potentiality for low delay and high error resilience to permit an easy extension to H.26x.

- To maximize the interoperability and applicability of the future HDTV system throughout the Federal government, it is recommended that the selected system be compatible with progressive scan and that packetized transmission be employed. This will maximize compatibility with computerized systems and future ATM networks.

2.0 STATUS OF HDTV STANDARDS

2.1 HDTV STANDARDIZATION IN THE U.S.

In the United States, the term "Advanced Television" (ATV), is a broad term which includes the three categories listed below.

IDTV - IMPROVED DEFINITION TELEVISION - The term Improved Definition Television refers to improvements to NTSC television which remain within the general parameters of NTSC emission standards and, as such, would require little television receiver modifications and may include improvements in encoding, filtering, ghost cancellation, and other parameters that may be transmitted and received as standard NTSC in a 4:3 aspect ratio.

EDTV - EXTENDED DEFINITION TELEVISION - The term Extended Definition Television refers to a number of different improvements that modify NTSC emissions but that are NTSC receiver-compatible (as either standard 4:3 or "letter-box" format). These changes may include one or more of the following:

1. Wide aspect ratio.
2. Extended picture definition at a level less than twice the horizontal and vertical emitted resolution of standard NTSC.
3. Any applicable improvements of IDTV.

HDTV - HIGH DEFINITION TELEVISION - The term High Definition Television refers to television systems with approximately twice the horizontal and vertical emitted resolution of standard NTSC. HDTV systems are wide aspect ratio systems and may include applicable improvements from IDTV and EDTV. IDTV and EDTV feature compatibility with the existing NTSC system which automatically puts severe limitations on their achievable performance. HDTV eliminates this constraint. This report is concerned only with the standardization of HDTV and its implementations for the Federal Government.

The purpose of this section is to provide an overview of the HDTV standardization work in the U.S. The description is divided into the four parts

listed below.

2.1.1	Federal Communication Commission (FCC)
2.1.2	Advanced Television Systems Committee (ATSC)
2.1.3	Society of Motion Picture and Television Engineers (SMPTE)
2.1.4	Proposed Digital HDTV Systems

2.1.1 Federal Communications Commission

In 1941, the Federal Communications Commission (FCC) made the first spectrum allocation and adopted the first technical standard for the transmission of black and white television in the U.S. [1]. Forty-six years later, in response to a joint petition filed by the Association of Maximum Service Telecasters, Inc. and 57 other broadcast organizations and companies, the Commission began a formal proceeding to replace that standard. On July 16, 1987, the Commission adopted a Notice of Inquiry in the matter of Advanced Television (ATV) Systems and their impact on the existing television broadcast service.

This notice went beyond the basic proposal to accommodate a new technology and raised the idea of delivering ATV by over-the-air terrestrial means, that is, by individual broadcast stations. This was a bold challenge considering that ATV research had concentrated on satellite and satellite-cable delivery systems, and many parties did not believe that terrestrial broadcast of high definition pictures was practical. Nonetheless, in order to make the benefits of an ATV service available to all viewers, the Commission was of the view that existing television licensees should be given the opportunity to provide that service.

Following the issuance of the Notice, the Commission established the Advisory Committee on ATV Service (ACATS). This group operates under a broad charter for the purpose of advising the Commission on the facts and circumstances regarding ATV systems, and recommending policies, standards, and regulations that would facilitate the orderly and timely introduction of ATV services in the U.S. In its activities, it may consider all technical, economic, legal, and regulatory issues.

The organizational structure of the ACATS is illustrated in Figure 2.1. The Planning Subcommittee has the task of defining the desirable characteristics of an ATV service and recommending planning factors for the establishment of that

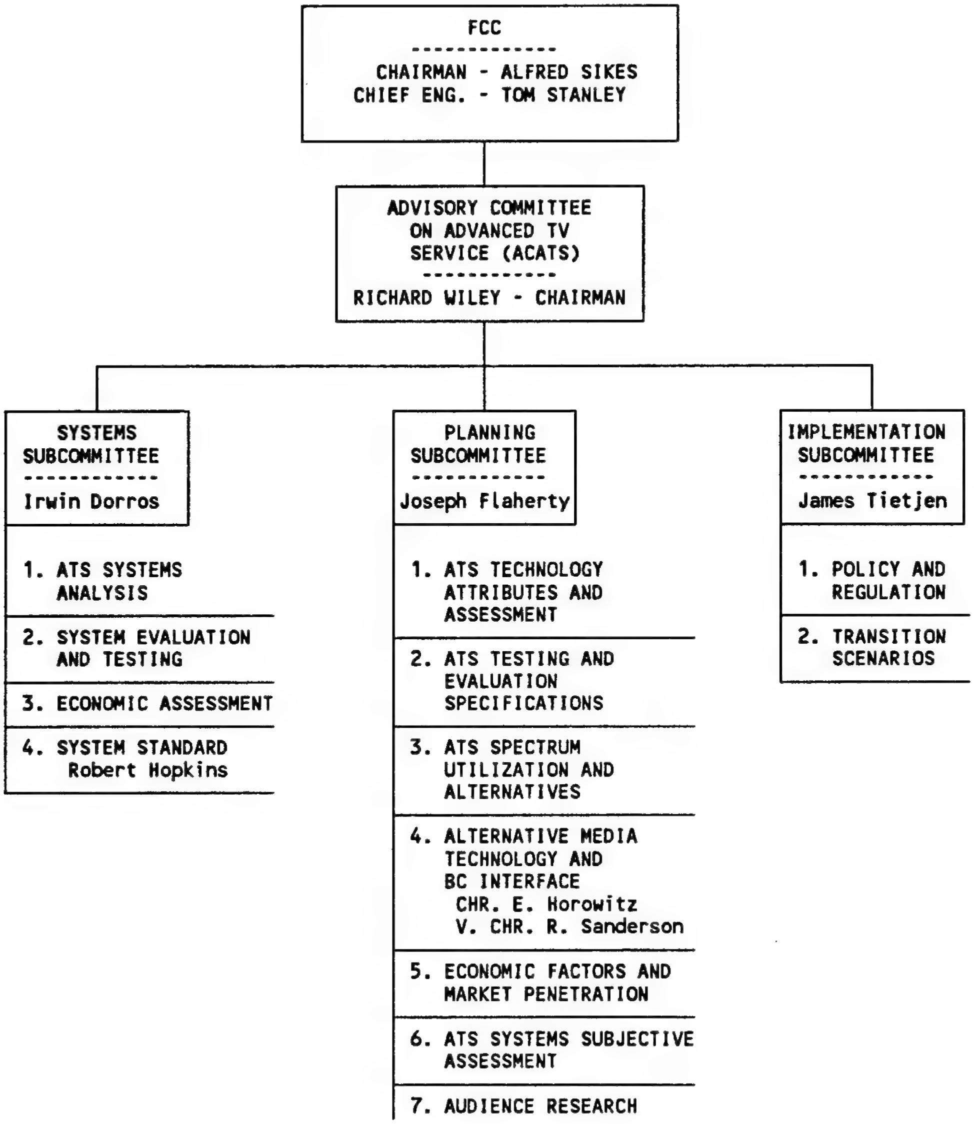


Figure 2.1
ACATS Organizational Structure

service. The Systems Subcommittee is hardware oriented; its task is to evaluate ATV systems under development, recommend a system (or specify the design of a system), and advise on the appropriate technical standards and spectrum requirements. Working Party 4 (Alternative Media Technology) is particularly important to the telecommunications community because its function is to insure that "inoperability" and "extensibility" are properly considered in the selection of

the HDTV standard.

The Implementation subcommittee deals with transition schemes and provides advice on official policies and regulations needed to implement an ATV service. At the present time, the members are concentrating on spectrum and implementation issues and the evaluation of five proposed ATV systems at the ATV Test Center. This test center is a facility established and funded by major elements of the television industry for the purpose of thoroughly measuring the attributes and performance of potential ATV systems under normal and adverse conditions. The results of the tests will be given to the Systems Subcommittee and will eventually be incorporated in the final report of the Advisory Committee to the Commission.

In a report adopted on August 24, 1990, the Commission made a critical technical decision and established a rough timetable for the completion of this proceeding and adoption of an ATV standard. The Commission stated that it intends to select a simulcast High Definition Television (HDTV) system for its ATV standard. *Simulcast* is a contraction of *simultaneous broadcast* and normally means the broadcast of one program over two channels in the same area at the same time. In this proceeding, simulcast has come to mean an independent ATV signal that can produce an advanced picture and be broadcast simultaneously with conventional television signals in the same area without causing interference.

Consistent with this position, the Commission also said that it will give no further consideration to systems that use additional spectrum for special signals to augment conventional television transmissions. That approach relies on ATV receivers to produce an advanced, high-quality picture by combining the separate conventional and augmentation signals. Such a system would be less spectrum efficient and more difficult to implement than the independent, simulcast approach.

Finally, the Commission urged the Advisory Committee to complete its work and submit a final report with recommendations by early 1993. The FCC is adhering to that schedule. Consequently, most of the work by the ACATS committee in 1992 was focused on the evaluation of the five proposed HDTV systems which are summarized below. The table includes the dates when tests were performed in 1992 by the Advanced Television Test Center (ATTC), CableLabs, and the ATEL laboratory in Canada.

PROPONENT	NAME	ATTC/CableLabs/ATEL TEST DATES*
NHK	Narrow Muse	9/30 - 2/3
General Instrument, MIT	DigiCipher	12/10 - 3/31
AT&T, Zenith	Digital Spectrum Compatible (DSC)	3/4 - 6/5
NBC, Philips, Sarnoff, Thomson	Advanced Digital Television	6/15 - 8/14
MIT, General Instrument	Channel Compatible DigiCipher (CCDC)	7/24 - 10/8

* All completion dates in 1992.

The first system transmits the signal in an analog format while the last four are all-digital. It is almost certain that one of the all-digital systems, or some hybrid combination thereof, will be selected. The four all-digital systems are sufficiently important that they are briefly described in Section 2.1.4.

Key events during 1992 which contributed to the evaluation of the five proposed HDTV systems are summarized below.

7/15 — PS-WP/4 meeting defines requirements for proponents at 9/23 Interoperability Review.

9/23-25 — PS-WP/4 Interoperability Review in Arlington, VA

+ Submission of written reports by the proponents describing their interoperability, extensibility, and scope of services/features.
+ Oral presentation of the reports to PS-WP/4 by the proponents.

October/ November	Formal Submission to ACATS of "Proposed Improvements" to the proposed HDTV systems by proponents. It is anticipated that some of these improvements will be tested in 1993.
December 10	Final Meeting of ACATS/PS/WP-4. Edited and approved the three key output documents. + Assessment of the Interoperability of the HDTV Proposals (Appendix A). + Analysis of Interoperability Evaluation by Proponent and Review Board (Appendix B). + PS-WP/4 Final Report (Appendix C).

Laboratory testing of the proposed systems is essentially complete and the following events are scheduled in the near future.

February 8-12, 1993	Meeting of the ACATS Special Panel to complete the evaluation of test data and make a recommendation to ACATS.
February 24, 1993	ACATS committee will make its recommendation to the FCC.
1993	Likely testing of improvements to one or more of the proposed systems (3 weeks each).
1993	Field test of the selected system.

An overview of key conclusions from the standardization work by the ACATS and PS/WP4 for 1992 as viewed from the Federal telecommunication's perspective is provided below.

- There are particular characteristics of the proposed HDTV systems which are advantageous from the Federal telecommunication's perspective.
  - + Progressive scan is advantageous because artifacts are

minimized, compression is maximized, and the highest level of compatibility with computers is achieved.

+ Packetized transmission with two levels of priority is advantageous because this is highly compatible with ATM telecommunication networks.
+ Square pixels are desirable because they maximize compatibility with computer systems which manipulate imagery such as in a windows environment.

- All of the digital HDTV proposals have the potential for a high level of interoperability, extensibility, and scalability. The Sarnoff system is probably most interoperable because...

  + the coding algorithm is MPEG/H.261 based,
  + the data is transmitted in packets which is well adapted to ATM,
  + the packets have two levels of priority which is well adapted to ATM, and
  + Headers/Descriptors are employed to provide extensibility and flexibility.

- One unique characteristic of the AT&T/Zenith DSC system is that the DCT coefficients are addressed using vector quantization. This is different from MPEG and therefore reduces interoperability with H.261.

- The AT&T/Zenith DSC and DigiCipher ATVA employ progressive scanning and square pixels and are therefore highly interoperable with the computer medium.

- The ACATS PS-WP/4 Final Report (Appendix C) properly highlights the need for headers/descriptors, square pixels, packetized transmission, and progressive scan.

2.1.2 Advanced Television Systems Committee

The Advanced Television Systems Committee (ATSC) was formed by the NAB, NTCA, IEEE, EIA, and SMPTE to further the work on HDTV. It is a free standing committee, and the organization is shown in Figure 2.2.

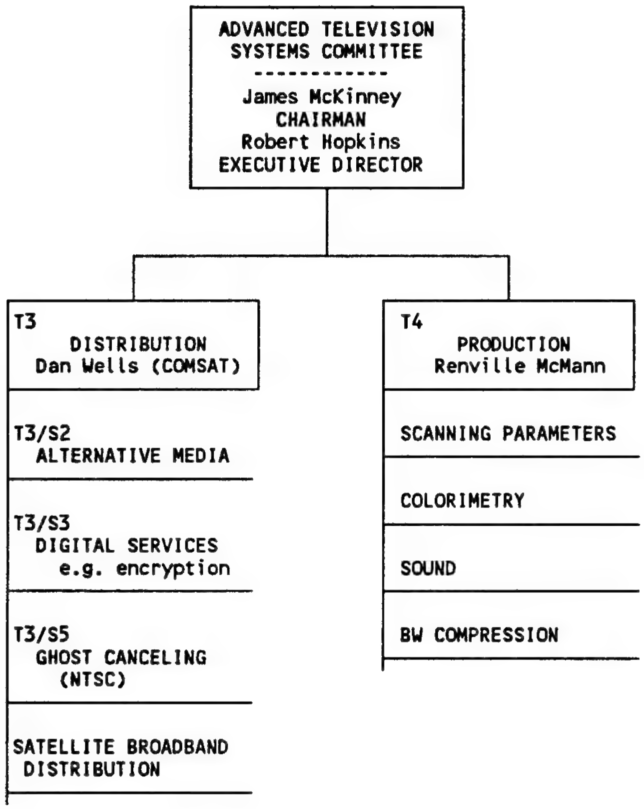


Figure 2.2
ATSC Organizational Structure

2.1.3 SMPTE

The SMPTE organization has been a major contributor to the work in HDTV standards in recent years. One major contribution has been the development of the 240M standard (1125 lines, 60 fields/sec.) for an analog HDTV signal which is used extensively in TV production studios in the US and elsewhere. The key

parameters for this standard are listed in Table 2-1.

Table 2-1
240M Key Parameters

IMAGE PARAMETERS	VALUE
TOTAL NUMBER LINES	1125
NUMBER ACTIVE LINES	1035
FIELD RATE	60
INTERLACE	2:1
ASPECT RATIO	16:9
BANDWIDTH (MHz):	
LUMINANCE	Y = 30
COLOR DIFFERENCE #1	Pr = 15
COLOR DIFFERENCE #2	Pb = 15

Work is underway on the development of a digital version of 240M. In the digital system there will be 1920 samples/line.

In 1992, SMPTE was actively involved in the two projects listed below which are closely related to the HDTV standardization process.

- Universal Header/Descriptor for Imagery
- Digital Image Architecture

Universal Header/Descriptor for Imagery

A SMPTE group was formed in June, 1991 to investigate the feasibility of a universal header/descriptor for imagery. The purpose of the header and descriptor is provided below.

- Header
    - + header is universally self identifying
    - + header determines how to interpret the payload
    - + minimal header (kernel) is identifier and packet length
    - + a packet would typically be a frame or group of frames

- Descriptor
    - + descriptor provides application oriented information
    - + e.g., picture size, frame rate, filters; production information; copyright; scrambling; etc.
    - + expert groups establish descriptors to address their needs

The committee has completed its work and the results are summarized in a report provided in Appendix D. Since the work is very promising, the SMPTE has formed a working group to develop a header/descriptor standard.

Digital Image Architecture

In April of 1991, SMPTE established a Task Force on Digital Image Architecture which "was charged with developing and proposing a structure for a hierarchy of digital image standards that would facilitate interoperation of image systems. The major objective was to establish the basis for image systems that are open, scalable and extensible, thus meeting the perceived needs for image communications in the environment likely to exist as computers, television and communications converge, enabled by pervasive digital technology.

The Task Force, formed from representatives of the affected industries and applications, has examined the issues, setting out those that are believed critical at this time, and has modelled, for discussion, further refinement and testing, one possible approach that meets the basic requirements. It has also produced extensive tutorial information concerning the matters under consideration.

The key concepts of the approach are that the conditions for image systems are:

- Open - the modules and interfaces forming the architecture are fully defined and in the public domain.

- Interoperable - images and related equipment may move freely across application and industry boundaries.

Such systems would be based on a hierarchy that is:

- Scalable - supports a wide range of image capabilities,
- Extensible - future-proof to the extent possible,
- Compatible - supports existing television practices and standards when possible.

The Task Force submitted a Final Report (Appendix E) to SMPTE in late 1992. It has been adopted, and a new group has been established to continue the work toward the development of a standard.

The report states the main objectives of the Task Force activity were:

- To establish the fundamental properties of image systems,
- To examine the technological trends with a view to a prediction of future capabilities,
- To arrive at architectural guidelines that will achieve the objectives of interoperability, scalability and extensibility.

The report details critical issues in the development of a suitable image architecture such as:

- The selection of a family of image acquisition rates and related display refresh rates based on a progression that permits display refresh at integer multiples of the acquisition rate. Backward compatibility , if required, to the image acquisition rates currently in use (24, 50, 59.94, and 60) should be accommodated in the design of the standard modules which will interface these existing systems with the digital image architecture,
- Use of a square sampling grid as a simple and effective common expression of images,
- Selection of analysis and coding schemes for color and luminance that would allow useful and effective scaling of this image data while

maintaining a high level of interoperability,

- Coherent sampling of the image based on the use of progressive scanning techniques,
- Use of headers/descriptors to identify the content and conform to the characteristics of the data stream.

2.1.4 Proposed Digital HDTV Systems

The purpose of this section is to briefly describe the four all-digital HDTV systems which have been proposed to the FCC. Much of the information used in these descriptions was taken from documents prepared by the proponents. Highlights of the four systems are summarized in Table 2-2.

2.1.4.1 Advanced Digital Television (ADTV)

Advanced Digital Television (ADTV) is the digital high-definition television (HDTV) system developed by the Advanced Television Research Consortium (ATRC). Figure 2.3 illustrates an architectural view of ADTV, which has been designed as a layered system. Four principal layers are shown: the compression layer, the data prioritization layer, the transport layer, and the transmission layer. The compression layer performs the tasks of video pre/post-processing and MPEG compression/decompression. The prioritization layer has the tasks of assigning priorities to video data at the encoder, and recombining the data elements of different priority into coherent data streams for decompression at the decoder. The data transport layer is responsible for service-independent data multiplexing, cell formatting, error detection, error correction, as well as service-specific logical error recovery. The transmission layer performs the tasks of modulation, channel equalization, and frequency translation.

Video compression in ADTV is based on the MPEG standard. While MPEG compression provides good image quality in the bit-rate regime of interest (i.e., 15-20 Mbps for HDTV resolution images), the particular robustness, co-channel immunity and service flexibility requirements of the broadcast application required substantial augmentation. The approach developed by the ATRC is called "MPEG++", in which an MPEG-compatible video representation is encapsulated with additional broadcast-specific prioritization and transport layers. A high degree

Table 2-2
Attributes, Characteristics, and Processes of
Digital HDTV Terrestrial Broadcasting Systems

System Name	DigiCipher	DSC-HDTV (Digital Spectrum Compatible)	ADTV (Advanced Digital Television)	CCDC-HDTV (Channel Compatible DigiCipher)
Proponent	General Instrument, MIT	AT&T, Zenith	NBC, Philips, Sarnoff, Thomson	MIT, General Instrument
Lines per Frame	1,050	787/788	1,050	787/788
Frames per Second	29.97	59.94	29.97	59.94
Interlace	2:1	1:1	2:1	1:1
Horizontal Scan Rate	31.469kHz	47.203kHz	31.469kHz	47.203kHz
Aspect Ratio	16:9	16:9	16:9	16:9
Active Video Pixels	1,408H x 960V (luma) 352H x 480V (chroma)	1,280H x 720V (luma) 640H x 360V (chroma)	1,440H x 960V (luma) 720H x 480V (chroma)	1,280H x 720V (luma) 640H x 360V (chroma)
Pixel Aspect Ratio	33:40	1:1	27:32	1:1
Bandwidth	21.5MHz (luma) 5.4MHz (chroma)	34MHz (luma) 17MHz (chroma)	24.5MHz (luma) 12.25MHz (chroma)	34MHz (luma) 17MHz (chroma)
Colorimetry	SMPTE 240M	SMPTE 240M	SMPTE 240M	SMPTE 240M (approx.)
Video Compression Algorithm	Motion-compensated DCT coding	Motion-compensated transform coding (DCT & VQ)	Motion-compensated DCT coding (MPEG-based)	Motion-compensated DCT coding
Block Size	8 x 8	8 x 8	8 x 8	8 x 8
Sampling Frequency	53.65MHz	75.3MHz	54MHz	75.5MHz
Audio Bandwidth	20kHz	20kHz	23kHz	24kHz
Audio Compression Algorithm	AC-2	AC-2	MUSICAM	MIT-AC
Audio Sampling Frequency	48kHz	47.203kHz	48kHz	48kHz
Dynamic Range	85dB	96dB	96dB	100db
Number of Audio Channels	4	4	4	4

System Name	DigiCipher	DSC-HDTV (Digital Spectrum Compatible)	ADTV (Advanced Digital Television)	CCDC-HDTV (Channel Compatible DigiCipher)
Video Data Rate	12.59Mbits/s (16 QAM) 17.49Mbits/s (32 QAM)	Automatically varies from 8.6 to 17.1Mbits/s	17.73Mbits/s	13.6 (16 QAM) 18.88 (32 QAM)
Audio Data Rate	0.503Mbits/s	0.5Mbits/s	0.512Mbits/s (nominal)	0.755Mbits/s
Control Data	126kbits/s	40kbits/s (spare)	40kbits/s (data)	126kbits/s
Ancillary	126kbits/s	413kbits/s	256kbits/s (nominal)	126kbits/s
Sync	N/A	292 to 544kbits/s	N/A	N/A
Total Data	19.51Mbits/s (16QAM) 24.39Mbits/s (32QAM)	11.1 to 21.0Mbits/s	24Mbits/s (nominal)	21.5Mbits/s (16QAM) 26.43Mbits/s (32QAM)
Error Correction Overhead	6.17Mbits/s	1.3 to 2.4Mbits/s	5.5Mbits/s	6.54Mbits/s
RF Modulation (Terrestrial)	16QAM or 32QAM	2-level and 4-level VSB	Spectrally shaped QAM	16QAM or 32QAM
3dB Bandwidth (Terrestrial)	4.88MHz	5.38MHz	5.28MHz	5.28MHz
C/N Threshold (Terrestrial)	12.5dB (16QAM) 16.5dB (32QAM)	16dB (4-level data) 10dB (2-level data)	16.1dB (High priority channel) 11.1dB (Standard Priority Level)	11.7 (16QAM) 15.7 (32QAM)
Channel Equalization (Ghost Cancelling)	-2 to +24 ms (multiple ghosts)	-2 to +20 ms (multiple ghosts)	16 ms (may be extended to 40 ms) -4 to +4 ms (prototype hardware)	-2 to +24 ms (multiple ghosts)

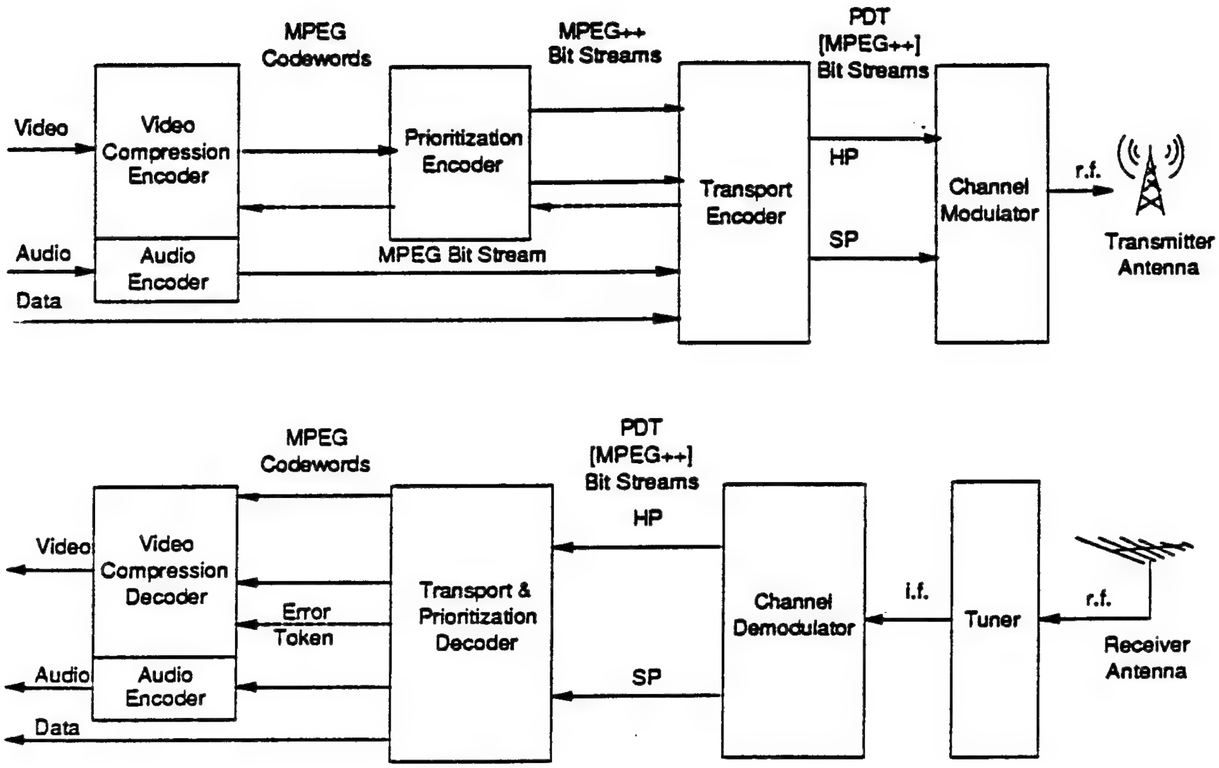


Figure 2.3
ADTV System High-Level Block Diagram

of robustness in the presence of channel impairments is achieved by first prioritizing MPEG encoded video data into High-Priority and Standard-Priority data streams, each of which is packaged into a rugged (and flexible) fixed-length "cell relay" type data format.

The Prioritized Data Transport format used in ADTV provides for reliable transport and synchronization of MPEG++ data over a two-tier physical transmission system. It uses a Reed-Solomon error correction code to fully correct the bit errors that can occur under moderate noise and interference conditions. Data interleaving ensures that even burst errors (which can be caused by impulse noise) can be fully corrected. The Prioritized Data Transport format also provides error detection capability, coupled with several transport features that support rapid video decoder recovery after uncorrectable channel errors. Such events will inevitably occur in a terrestrial broadcast channel during severely impaired transmission, and must be handled gracefully by a receiver.

The high- and standard-priority MPEG++ bit-streams are sent over the simulcast channel with appropriate transmission priority using a two-tier modulation technique which delivers approximately 20% of the data (i.e., the high priority information) at a substantially improved reliability level. In addition to providing two-tier transmission matched to the video prioritization process, ADTV's two-carrier SS-QAM approach mitigates the effect of interference from NTSC co-channel, thus permitting the low-power operation necessary for simulcast. Finally, reliable demodulation of the low-power ADTV signal is achieved without sacrificing bandwidth efficiency through the use of an appropriate high-rate trellis code within the QAM modem. At a systems level, MPEG++ compression and Prioritized Data Transport, together with ATRC's two-tier SS-QAM modulation technique, provide the combination of image quality, robustness, co-channel immunity and flexibility necessary for the digital simulcast application.

2.1.4.2 Channel-Compatible DigiCipher (CCDC)

The Channel-Compatible DigiCipher (CCDC) HDTV system is proposed by the Massachusetts Institute of Technology and General Instrument Corporation. A high resolution progressively scanned baseline video signal of 1280 x 720 picture elements, 59.94 fps, and 16:9 aspect ratio, can be transmitted within a single 6 MHz channel. The CCDC system is source-adaptive. The system can recognize

and adapt itself to the particular characteristics of the source format so that the highest quality video can be reconstructed. The system, for example, accommodates different frame rates. The system automatically optimizes both the encoding and decoding processes so that the highest quality video can be obtained for each source format. The system will support, for example, a wide range of source adaptivity including different frame rates, resolutions, aspect ratios, progressive/interlaced scanning, and the black-and-white/color characteristics of the video.

The system is also scalable, in that performance subsets can be extracted from the CCDC HDTV signal. With proper design of the receiver, the received video may be displayed at multiple resolutions with modest amounts of signal processing. As an example, lowpass filtering and subsampling the signal is an effective method to obtain a lower spatial resolution image. In addition, row/column doubling methods can be used to increase the spatial resolution.

A block diagram of the CCDC HDTV system is shown in Figure 2.4. The system consists of three parts: video coding, audio coding and transmission. By coding, we mean to remove redundancy in the information source and to transmit only the essential information.

A transmission or digital communications systems is required to link the digital outputs of the various source coders to an analog RF channel. In a good design, the transmission system isolates the source coder from the channel and the two problems of representing the HDTV signal efficiently and transporting it from broadcaster to user become largely independent. The transmission system for a given medium (terrestrial broadcast, cable, satellite, videotape, etc.) must account for imperfections and noise in the channel and deliver a stream of bits as reliably as possible.

In the case of video source coding, there is redundancy along all three dimensions of the signal. Temporal redundancy is removed by a motion-compensated predictive coding scheme in which the motion between the previous and current frames is estimated, a predicted current frame is computed based on the previous frame and the extrapolated motion, and only the new information is sent to the spatial coder.

Spatial redundancy arises from objects which have extent and texture; these introduce correlations between neighboring pixels. The correlations are removed using the Discrete Cosine Transform (DCT). The DCT coefficients are then

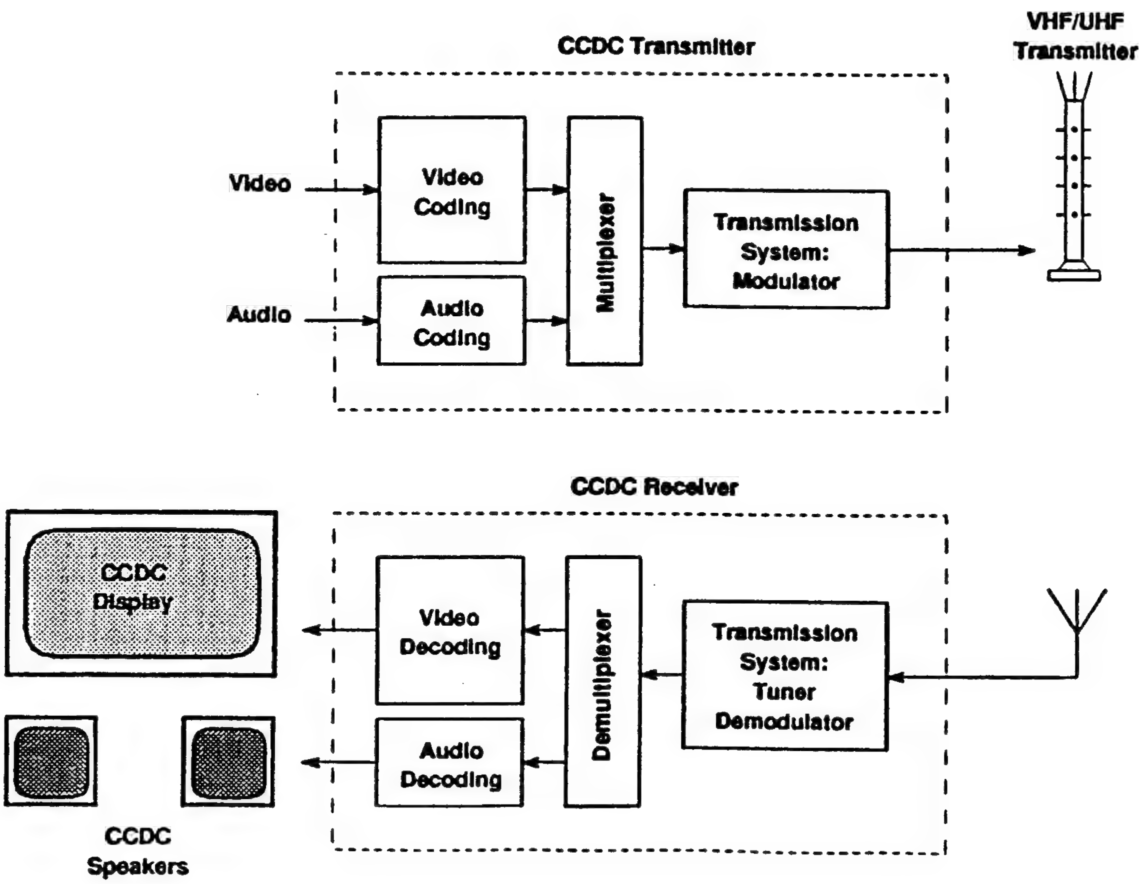


Figure 2.4
CCDC High-Level Block Diagram

weighted to reflect the relative importance of different frequency components, and coefficients which are perceptually important are quantized and entropy coded.

In audio coding, the input signal is divided into short overlapping segments, then windowed and transformed. These transform coefficients are then divided approximately into the critical bands of the human auditory system. Psychoacoustic models are used to eliminate inaudible information in the signal spectrum. Spectral envelope information is encoded using a fixed bit allocation strategy, and this envelope is used for dynamic bit allocation for the transform coefficients.

The encoded video data, audio data and auxiliary data are multiplexed into a single bit stream for the transmission system. Block and trellis encoding are used to reintroduce some redundancy to combat channel noise, and the resulting digital symbols are transformed into analog waveforms through quadrature modulation. The system can operate at 32 QAM (preferred mode) or 16 QAM, depending on the desired noise threshold/coverage area.

2.1.4.3 Digital Spectrum Compatible (DSC) HDTV System

Zenith and AT&T have developed the all-digital DSC high-definition television (HDTV) simulcast system. The system's 787.5-line progressive scanning format eliminates artifacts of interlaced systems, provides full motion rendition (critical for sports and other fast-action programming) and promotes HDTV compatibility with current and future computer and digital communications technologies. Progressive scanning plus square pixels and unique compression technology make the DSC-HDTV system well suited for interconnectivity, extensibility, scalability and other computer related considerations.

The DSC-HDTV system employs a four-level vestigial sideband (4-VSB) modulation technique which is complemented by a two-level digital data system (2-VSB) expanding the service area of the Digital Spectrum Compatible signal. The resulting *bi-rate coding system* identifies and selects the most important picture information on a scene-by-scene basis and automatically transmits that data in a two-level (binary) mode. The remainder of the picture information is transmitted as 4-level digital data. Two-level digital coding makes the system far more tolerant of noise and other interference at greater distances from the transmitter. This allows extended reception of the signal beyond the traditional NTSC service area and

eliminates the so-called "cliff effect", or complete and abrupt loss of picture, associated with some other all-digital approaches.

Picture Format

Digital Spectrum-Compatible High-Definition Television (DSC-HDTV) images are progressively scanned at 787.5 lines/frame, 59.94 frames/second. The aspect ratio is 16:9 and the horizontal line rate is 47.203 kHz, three times that of NTSC. The nominal video baseband signal bandwidth is 34 KHz. (With a Kell factor of 0.9 for sampling in the horizontal direction and an approximate sampling rate of 75.3 MHz, the nominal video bandwidth equals 0.9 x 75.3/2 = 33.9 MHz.)

Transmitter and receiver signal processing are performed on square pixels in a 720 line by 1280 pixel array. In the studio, an additional guard band of pixels at all four edges is provided to allow for transient effects of processing, analog rise times, production related edge effects and timing tolerances.

Square pixels are chosen to facilitate computer interface and special effects processing. the particular numbers offer easy conversion to/from NTSC for simulcast purposes. Conversion to 525 line CCIR 601 requires only a 4:3 interpolation horizontally and 3:2 vertically. In addition, the simple relationship to NTSC provides economical means for designing dual purpose HDTV/NTSC receivers.

Progressive scanning has advantages in video compression and display of motion without line pairing or resolution loss. The 1280 by 720 format is simply related to Common Image format by a linear factor of 2:3 and is easily extensible to higher-line-number progressive formats when they become practicable. The data rate generated by this format is within the capability of current commercial high definition tape recorder technology.

Video Coding

The Video Encoder accepts RGB signals with SMPTE 240M colorimetry and transfer function. The video encoder is shown in Figure 2.5. Motion from frame to frame is estimated using a hierarchical block-matching Motion Estimator. The Motion Estimator produces motion vectors, which are compressed and sent to the output buffer for transmission. The output buffer has an output rate varying

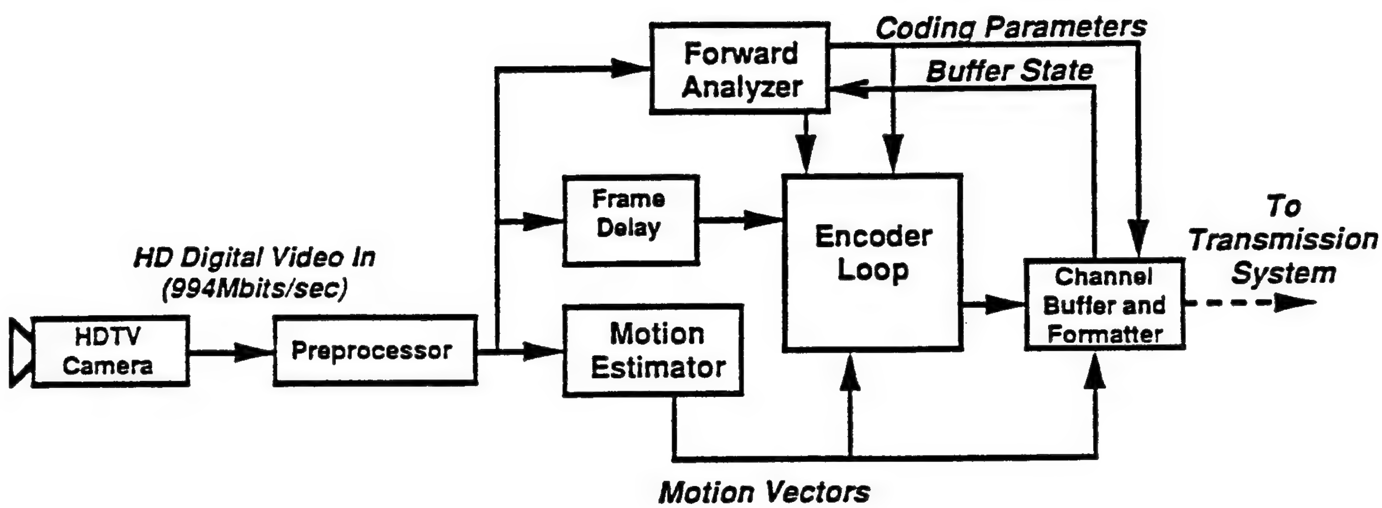


Figure 2.5
Video Encoder

between 9 and 17 Mb/s and has a varying input rate that depends on the image content. The buffer history is used to control the coding parameters so that the average input rate equals the average output rate. The feedback mechanism involves adjustment of the allowable distortion level, since increasing the distortion level (for a given image or image sequence) causes the Encoder to produce a lower output bit rate.

2.1.4.4 DigiCipher™ HDTV System

The DigiCipher™ HDTV system has been proposed by General Instrument for the American Television Alliance. Figure 2.6 shows the block diagram of the encoder. The digital video encoder accepts YUV inputs with 16:9 aspect ratio and 1050-line interlace (1050/2:1) at a 59.94 field rate. The YUV signals are obtained from analog RGB inputs by low pass filtering, A/D conversion, and an RGB-to-YUV matrix. The sampling frequency is 53.65 MHz for R,G, and B. The digital video encoder implements the compression algorithm and generates a video data stream. The digital audio encoder accepts four audio inputs and generates an audio date stream. The data/text processor accepts four data channels at 9600 baud and generates a data stream. The control channel processor interfaces with the control computer and generates control data stream.

The multiplexer combines the various data streams into one data stream at 18.22 Mbps. The FEC encoder adds error correction overhead bits and provide 24.39 Mbps of data to the 32-QAM modulator. The symbol rate of the 32-QAM signal is 4.88 MHz.

The DigiCipher™ HDTV system can also support 16-QAM transmission that provides lower system threshold with a slight penalty in picture quality. The lower system threshold can be used to improve the ATV coverage area and/or to reduce the station spacing. Through the use of a unique design, the DigiCipher™ system shares the same digital video/audio data processing, forward error correction, and QAM modulation/demodulation circuities between the two operating modes. Table 2-3 summarizes the key parameters of the DigiCipher™ HDTV system.

The video compression process can be broken down into the following different subprocesses:

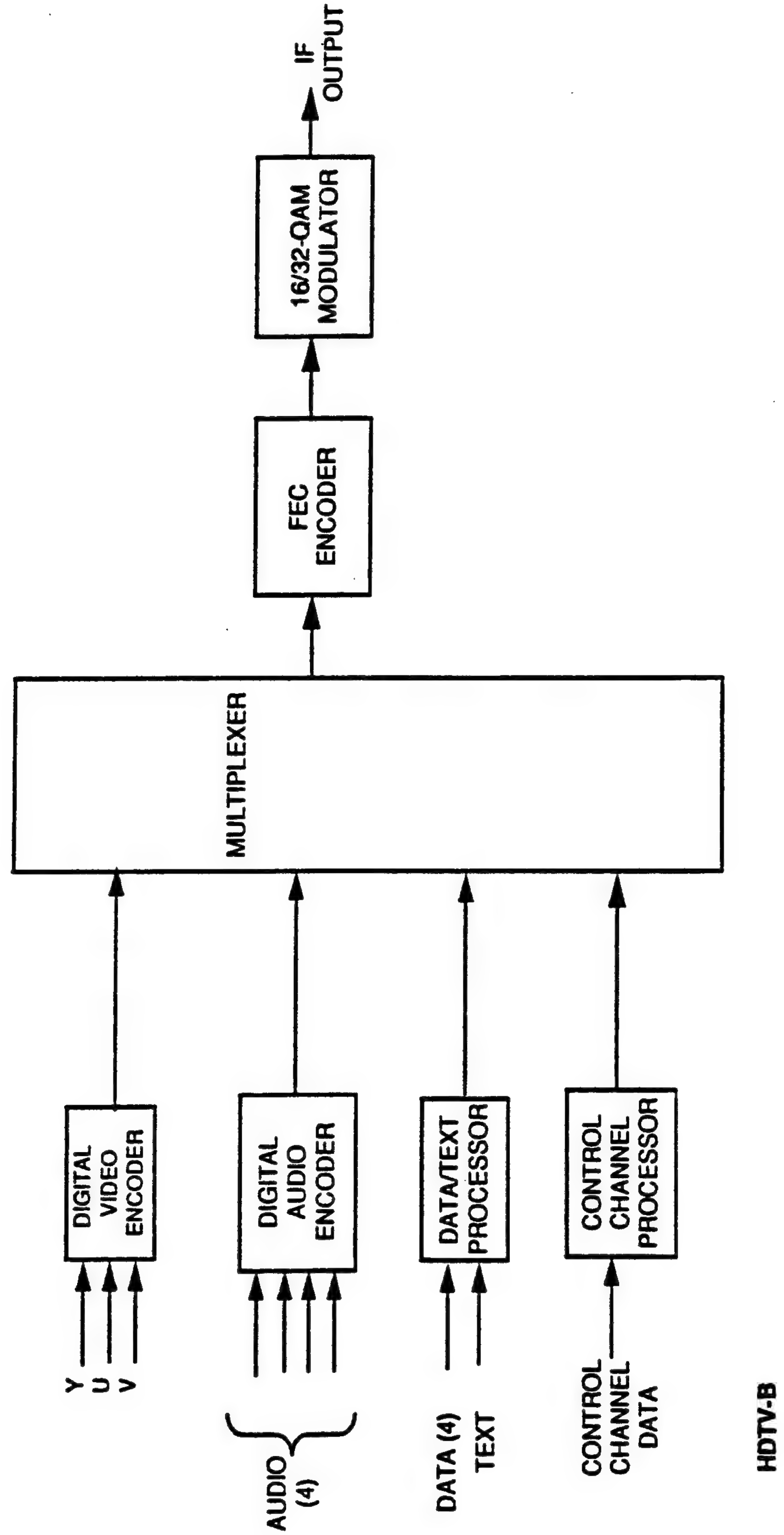


Figure 2.6
Encoder Block Diagram

Table 2-3
DigiCipher™ System Parameters

Operating Mode	16-QAM	32-QAM
VIDEO		
Raster Format	1050/2:1 Interlaced	1050/2:1 Interlaced
Aspect Ratio	16:9	16:9
Frame Rate	29.97 Hz	29.97 Hz
Bandwidth		
Luminance	21.5 MHz	21.5 MHz
Chrominance	5.4 MHz	5.4 MHz
Active Pixels		
Luminance	960(V) x 1408(H)	960(V) x 1408(H)
Chrominance	480(V) x 352(H)	480(V) x 352(H)
Horizontal Resolution		
Static	660 Lines per Picture Height	660 Lines per Picture Height
Dynamic	660 Lines per Picture Height	660 Lines per Picture Height
Sampling Frequency	53.65 MHz	53.65 MHz
Colorimetry	SMPTE 240M	SMPTE 240M
Horizontal Line Time		
Active	26.24 usec	26.24 usec
Blanking	5.54 usec	5.54 usec
AUDIO		
Number of Channels	4	4
Bandwidth	20 kHz	20 kHz
Sampling Frequency	47.2 kHz	47.2 kHz
Dynamic Range	90 dB	90 dB
DATA		
Video Data	**12.59 Mbps**	**17.47 Mbps**
Audio Data	503 kbps	503 kbps
Async Data and Text	126 kbps	126 kbps
Control Channel Data	126 kbps	126 kbps
Total Data Rate	**13.34 Mbps**	**18.22 Mbps**
TRANSMISSION		
FEC Data	6.17 Mbps	6.17 Mbps
DATA Transmission Rate	**19.51 Mbps**	**24.39 Mbps**
QAM Symbol Rate	4.88 MHz	4.88 MHz
Adaptive Equalizer Range	-2 to 24 usec	-2 to 24 usec
SYSTEM THRESHOLD		
Noise (C/N)	**12.5 dB**	**16.5 dB**
ATV Interference (C/I)	**12.0 dB**	**16.0 dB**
NTSC Interference (C/I)	**0.0 dB**	**5.0 dB**

NOTE: The differences in the two operating modes are emphasized in bold letters.

1. A/D Conversion and RGB-to-YUV Matrix
2. Chrominance Preprocessor
3. Discrete Cosine Transform (DCT)
4. Coefficient Quantization
5. Huffman (Variable Length) Coding
6. Motion Estimation and Compensation
7. Integration of Motion Compensation with Intraframe Coding
8. Adaptive Field/Frame Processing
9. Motion Picture Processing
10. Rate Buffer Control

A block diagram for the encoder video processing is shown in Figure 2.7.

2.2 INTERNATIONAL HDTV STANDARDS ACTIVITY

In 1992 there was great interest and activity in HDTV around the world. In Japan, the MUSE (Multiple Sub-Nyquit Encoding) system continues to be operational as an official service offered to the public. In Europe, the MAC (Multiplexed Analog Components) system continues to be evaluated on an experimental basis. Both of these systems employ analog transmission techniques as opposed to the all-digital approach likely to be selected by the U.S. During 1992, the world has begun to believe that the all-digital approach will be the ultimate winner and has stepped up their research, development, and testing work in the all-digital HDTV area.

During 1992, there was considerable activity by the CCIR toward the development of an international all-digital HDTV standard. Special Rapporteur Group II (User Requirements) prepared a key document entitled "Statement Concerning User Requirements for Source Coding and Multiplexing for Broadcast Applications" which is included in Appendix F. A few key statements from this document are provided below.

1. It is widely recognized today that broadcasting is in the middle of a revolution leading towards the "all-digital" broadcast system of the future.
2. It is appropriate at this time to draw attention to the great interest in

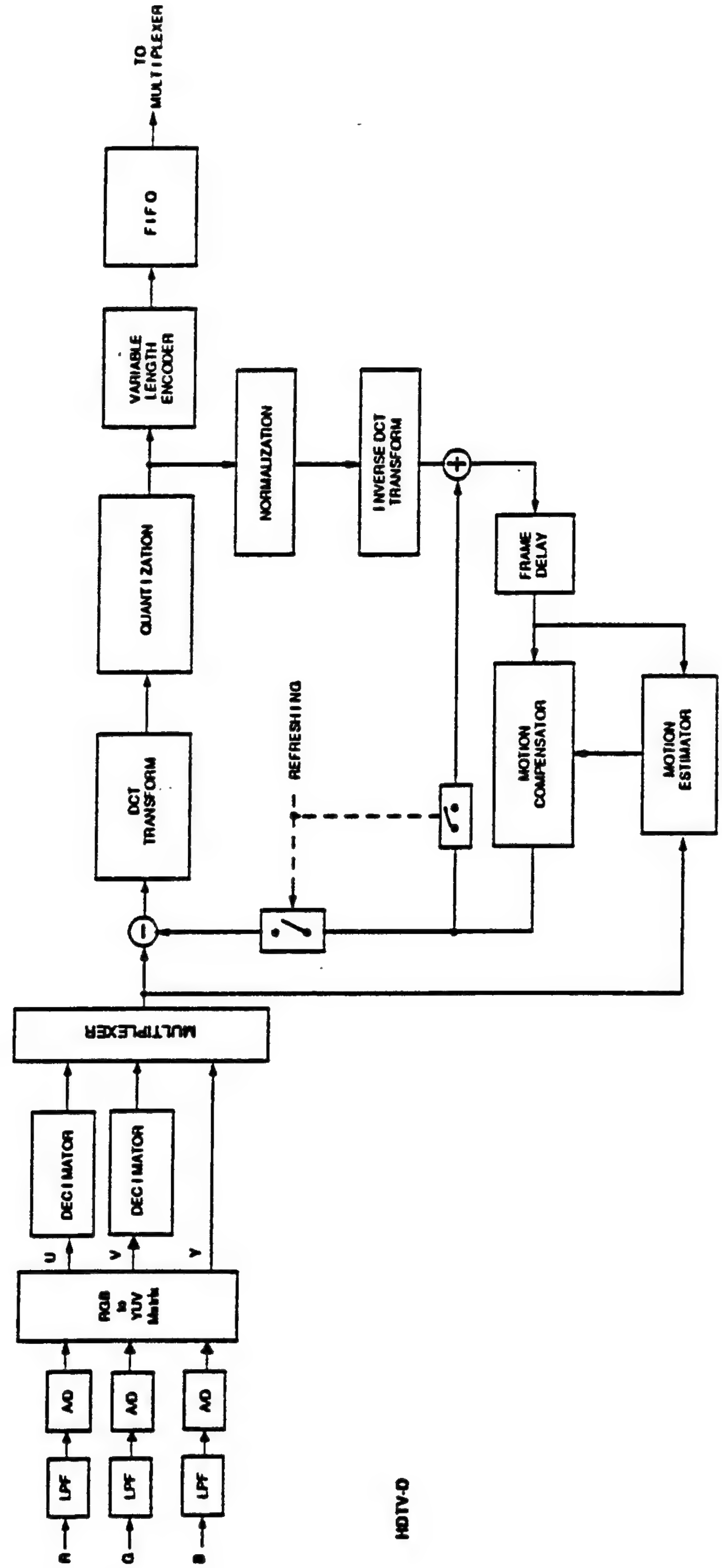


Figure 2.7
Digital Video Encoder Block Diagram

systems able to support hierarchical source and channel coding to provide HDTV, EDTV and CDTV (STV) transmission and receiver options for a broadcast signal with transmission bit rates in the range 20-30 Mbit/s.

3. In recognition of the rapid time-scale envisaged for standardization within MPEG, the Special Rapporteur Group is charged with the production of its first Report concerning the requirements for broadcasting and secondary distribution in September, 1992.

There are three Task Groups within the CCIR that are working on HDTV standards.

- **Task Group 11/1** - Developing a standard for HDTV production and computer processing.

- **Task Group 11/3** - Working on a digital standard for terrestrial broadcast. Stan Baron, of NBC, is the Chairman.

- **Task Group 11/4** - The purpose of this group is to harmonize HDTV work within, and outside the CCIR. This group met in Washington in October. To achieve harmony, the group advocates common source algorithms, hierarchical coding, and headers/descriptors. This concept is outlined in a technical contribution which is included in Appendix G.

3.0 STATUS OF TELECOMMUNICATION STANDARDS RELATED TO HDTV

There are four digital TV telecommunication standards that are closely related to HDTV. These are:

- H.261
- MPEG1
- H.26X
- MPEG2

H.261 was the first to be developed. It was designed for interactive teleconferencing at bit rates from 64 Kbps to 2 Mbps. MPEG1 built on many of the ideas of H.261, and added some new ones. Its primary focus was VHS quality video from digital storage media at about 1.5 Mbps, although a wide range of quality and bit rates are allowed. MPEG2 is an advanced version of MPEG1 with many additional features. It is designed for higher bit rates (4 and 9 Mbps were used in simulations) and very high quality CCIR 601 resolution. H.26X shares a common text with MPEG2, but selects parameters and modes of operation that are suitable for interactive communications. It is designed for use with BISDN networks, using ATM access and possibly variable bit rates.

The following sections describe these standards in some detail, and indicate their relationship to each other and to the proposed HDTV standards.

3.1 CCITT RECOMMENDATIONS

3.1.1 H.261

In December 1990, the CCITT finalized a set of five Recommendations (H.261, H.221, H.242, H.230, and H.320) which collectively define an audiovisual terminal to provide video teleconferencing (VTC) and video telephony (VT) services over the Integrated Services Data Network (ISDN). Since the basic building block of the ISDN is a basic channel (BRI) operating at 64 Kbps, the generic term "P x 64 Kbps" refers to operation of this terminal at integral values of P up to a maximum of 30.

H.261 is entitled "Video Codec for Audiovisual Services at P x 64 Kbps" and

fully defines the video codec function for the P x 64 audiovisual terminal including picture format, video coding algorithm and forward error control. Figure 3.1 is a functional block diagram of the H.261 video codec. The heart of the system is the source coder which compresses the incoming video signal by reducing redundancy inherent in the TV signal. The multiplexer combines the compressed data with various side information which indicates alternative modes of operation. A transmission buffer is employed to smooth the varying bit rate from the source encoder to adapt it for the fixed bit rate communication channel. A transmission coder includes functions such as forward error control to prepare the signal for the data link.

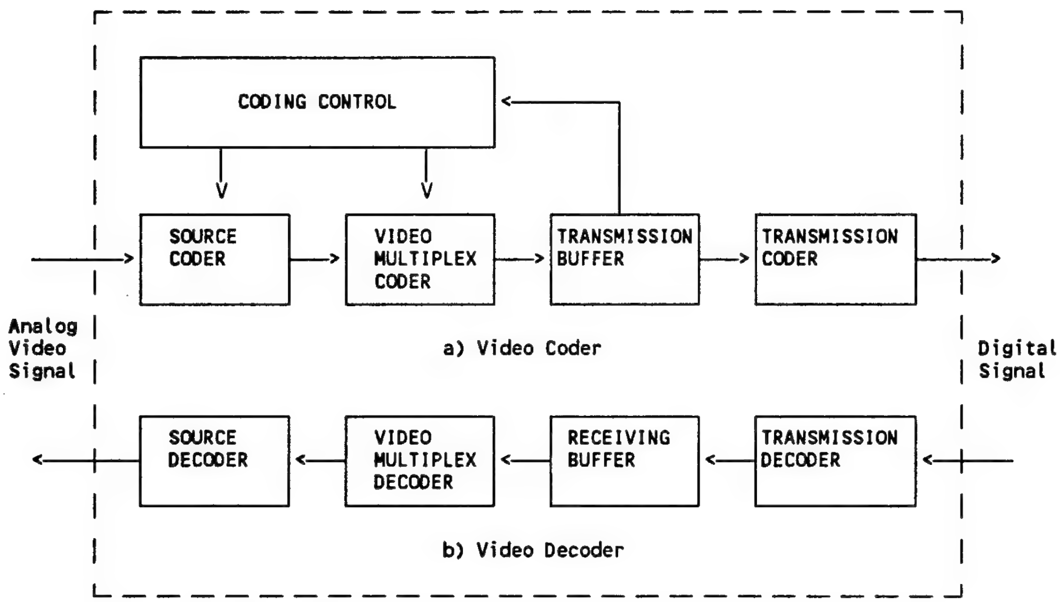


Figure 3.1
Block Diagram of the Video Codec

Since H.261 is a true international standard, there is a basic challenge to reconcile the incompatibility between European TV standards (PAL, SECAM) and those in most other areas of the world (NTSC). PAL and SECAM employ 625 lines and a 50 Hz field rate while NTSC has 525 lines and a 60 Hz field rate. This conflict was solved by adopting a Common Intermediate Format (CIF) and QCIF (Quarter CIF) as the picture structure which must be employed for any transmission adhering to H.261. The CIF and QCIF parameters are defined below.

	CIF	QCIF
Coded Pictures per second	29.97	(or integral submultiples)
Coded Luminance pixels per line	352	176
Coded Luminance lines per picture	288	144
Coded Color pixels per line	176	88
Coded Color lines per picture	144	72

The QCIF format, which employs half the CIF spatial resolution in both horizontal and vertical directions, is the mandatory H.261 format: full CIF is optional. It is anticipated that QCIF will be used for videophone applications where head-and-shoulders pictures are sent from desk to desk. Conversely, it is assumed that the full CIF format will be used for teleconferencing where several people must be viewed in a conference room.

Figure 3.2 is a functional block diagram outlining the H.261 source coder. Interframe prediction is first carried out in the pixel domain. The prediction errors are encoded by the Discrete Cosine Transform using blocks of 8 pels x 8 pels. The Transform coefficients are next quantized and fed to the multiplexer. Motion compensation is included in the prediction on an optional basis.

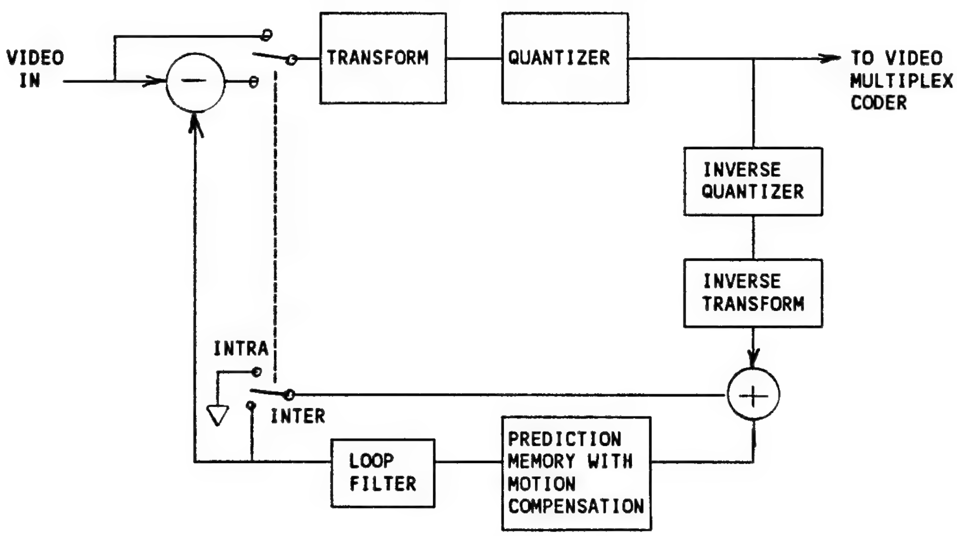


Figure 3.2 Source Coder

Picture Structure

In the encoding process, the picture is subdivided into macroblocks where each macroblock consists of a 2 x 2 array of DCT blocks (see Figure 3.3). The macroblock header defines the type of coding to be performed, possible motion vectors, and which blocks within the macroblock will actually be coded. There are two basic types of coding. In Intra coding, coding is performed without reference to previous pictures. This mode is relatively rare, but is required for forced updating, and every macroblock must occasionally be Intra coded to control the accumulation of inverse transform mismatch error. The more common coding type is Inter, in which only the difference between the previous picture and the current one is coded. Of course, for picture areas without motion, the macroblock does not have to be coded at all.

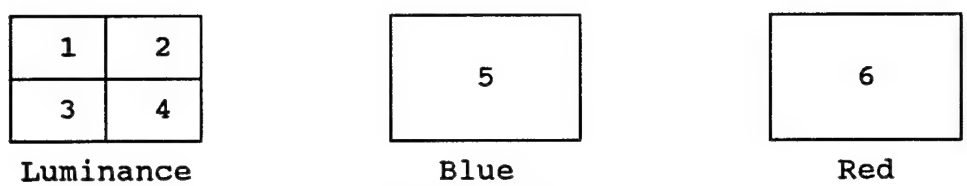


Figure 3.3
Arrangement of Blocks in a Macroblock

Example of Block Coding

Figure 3.4 shows a simple example of how each 8 x 8 block is coded. In this case, Intra coding is used, but the principle is the same for Inter coding. Figure 3.4a shows the original block to be coded. Without compression, this would take 8 bits to code each of the 64 pixels, or a total of 512 bits. First, the block is transformed, using the two-dimensional Discrete Cosine Transform (DCT), giving the coefficients of Figure 3.4b. Note that most of the energy is concentrated into the upper left-hand corner of the coefficient matrix. Next, the coefficients of Figure 3.4b are quantized with a step size of 6. (The first term {DC} always uses a step size of 8.) This produces the values of Figure 3.4c, which are much smaller in magnitude than the original coefficients and most of the coefficients become zero. The larger the step size, the smaller the values produced, resulting in more compression.

75	76	77	78	79	80	81	82
77	78	79	80	81	82	83	84
79	80	81	82	83	84	85	86
81	82	83	84	85	86	87	88
83	84	85	86	87	88	89	90
85	86	87	88	89	90	91	92
87	88	89	90	91	92	93	94
89	90	91	92	93	94	95	96

a) ORIGINAL BLOCK (8x8x8 = 512 BITS)

684	-19	-1	-2	0	-1	0	-1
-37	0	-1	0	0	0	0	-1
0	0	0	0	0	0	0	0
-4	-1	-1	-1	-1	0	-1	-1
0	0	0	0	0	0	0	0
-2	0	0	-1	0	-1	0	-1
0	0	0	0	-1	-1	-1	-1
-1	-1	-1	0	-1	0	-1	0

b) TRANSFORMED BLOCK COEFFICIENTS

86	-3	0	0	0	0	0	0
-6	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

c) QUANTIZED COEFFICIENT LEVELS

RUN	LEVEL	CODE
0	86	01010110
0	-3	001011
0	-6	001000011
	EOB	10

TOTAL CODE LENGTH = 25

d) COEFFICIENTS IN ZIG-ZAG ORDER AND VARIABLE LENGTH CODED

688	-21	0	0	0	0	0	0
-39	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

e) INVERSE QUANTIZED COEFFICIENTS

76	76	77	79	80	81	82	83
77	77	78	80	81	82	83	84
79	79	80	81	83	84	85	86
81	82	83	84	85	87	88	88
84	84	85	87	88	89	90	91
86	87	88	89	91	92	93	93
88	89	90	91	92	94	95	95
89	90	91	92	93	95	96	96

f) RECONSTITUTED BLOCK

Figure 3.4
Sample Block Coding

The coefficients are then reordered, using the Zig-Zag scanning order of Figure 3.5. All zero coefficients are replaced with a count of the number of zero's before each non-zero coefficient (RUN). Each combination of RUN and VALUE produces a Variable Length Code (VLC) that is sent to the decoder. The last non-zero VALUE is followed by an End of Block (EOB) code. The total number of bits used to describe the block is 25, a compression of 20:1.

3.1.2 H.26X

In 1990, CCITT Study Group XV established the "Experts Group for ATM Video Coding" to develop a Recommendation for a video codec for operation over future ATM BISDN networks. The group has met nine times and has worked very effectively to achieve its objective. The work of the Experts Group has been

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

Figure 3.5
Scanning Order in a Block

tightly coordinated with the ISO MPEG Group (ITC1/WG11) as they work on the MPEG2 standard. The integration of these activities has been accomplished by the same personnel meeting as an ATM Experts Group but also as members of the MPEG group. This process has proven to be successful because the basic approach to MPEG2 has the following desirable characteristics.

- There is a low delay mode where B frames are not employed,
- there is much attention given to the coding algorithm to insure a high degree of error resilience which occurs more in a telecommunication environment,
- the MPEG2 algorithm is an extension of the H.261/MPEG1 algorithm which makes it possible to achieve backward compatibility at a reasonable cost.

Another comment on the standardization of H.26X are provided below.

- H.261 has a rigidly defined picture format -- QCIF, FCIF. It is likely that H.26X will be less rigid -- i.e., a number of formats will be defined as options.

3.2 ISO STANDARDS

The International Standards Organization (ISO) develops standards in many areas, but is primarily concerned with computer rather than communications issues. It set up a working group in 1988 to develop a standard for storing and retrieving video sequences. This group is officially ISO-IEC JTC1/SC2/WG11, but is more commonly known as Moving Picture Experts Group (MPEG). Other related groups are Joint Photographic Experts Group (JPEG) and Joint Bi-level Image Experts Group (JBIG).

3.2.1 MPEG1

This is officially International Standard (IS) 11172, "Coding of moving picture and associated audio--for digital storage media at up to about 1.5 Mbit/s", Rev 2, March 27, 1992. It is the original standard developed by the MPEG committee, and is sometimes referred to as simply MPEG. The MPEG1 standard specifies the coded representation of video for digital storage media, and describes the decoding process. The representation supports normal speed forward playback, as well as special functions such as random access, fast play, fast reverse play, normal speed reverse playback, pause and still pictures. It is compatible with standard 525- and 625-line television formats, and it provides flexibility for use with personal computers and workstation displays.

MPEG1 is primarily intended for application to digital storage media supporting a continuous transfer rate up to about 1.5 Mbps, such as Compact Disk (CD), Digital Audio Tape (DAT), and magnetic hard disks. The storage media may be directly connected to the decoder, or via communications means such as busses, LANs, or telecommunications links. MPEG1 is intended for non-interlaced video formats having approximately 288 (240 for NTSC) lines of 352 pels and picture rates around 24 Hz to 30 Hz.

While MPEG1 was built to a large degree on the concepts of H.261, because of its different application a number of new techniques were introduced.

The primary difference is the use of I, P, and B frames. Rather than encode any macroblock in Intra or Inter mode, MPEG1 designates certain frames as I frames, to be coded completely in the Intra mode. This makes it possible to achieve random access wherever an I frame is used. Other frames are designated

as P frames, and are coded relative to a previous I frame or a previous P frame, primarily using Inter coding with motion compensation. So far, this result could have been achieved by means of appropriate use of the tools that H.261 provides. The real innovation is the use of B frames, which are interpolated between preceding and following I and P frames. Interpolation is performed by, for each macroblock, predicting pixels relative to the immediately preceding I or P frame, or the immediately following I or P frame, or both, using motion vectors. When both are used, the two predictions are averaged. Then the difference between the prediction and the actual frame is coded using DCT. The B frame itself is never used to predict another B frame. The use of B frames provides a significant amount of additional compression, particularly in cases where a moving object uncovers background, which cannot be predicted from a previous frame, but can be from a subsequent frame. A disadvantage of the B frame is that it involves a considerable coding and decoding delay, and therefore is not suitable for interactive applications.

Another innovation is the use of motion vectors that have half-pixel resolution, rather than the full-pixel resolution used in H.261. Because of the improved quality with half-pixel resolution, the use of loop filtering is not required.

The MPEG1 standard has been technically frozen since May 1991, and is about to be formally approved. A substantial number of applications are already on the market.

The importance of MPEG1 for HDTV is that it forms the basis for all of the proponent digital systems, particularly the ADTV system, which follows it very closely. However, all of the proponents use motion-compensated DCT coding.

3.2.2 MPEG2

The MPEG2 standard is derived from MPEG1. Its primary difference is that it provides higher resolution and higher quality at higher bit rates. The nominal resolution is CCIR 601, which for NTSC is 720 pixels per line and 480 active lines. The standard is useful over a wide range of bit rates, but during its development experiments were performed at 4 and 9 Mbps. At 9 Mbps, the quality of the decoded images is such that it is usually very difficult to distinguish them from the original digitized images.

In addition to the MPEG1 requirements, there are additional requirements in

the areas of compatibility with MPEG1 and H.261, error and cell-loss resilience, extensibility, scalable decoding, channel hopping, bit stream editing, and variable bit rate transmission.

A major difficulty with achieving higher vertical resolution is that interlaced formats must be dealt with. For MPEG1 the vertical resolution (240 lines) was such that only one field from an interlaced frame had to be used. For MPEG2 (480 lines) both fields must be used. Special prediction methods are required to make the best trade-off between using the pixel that is closest in time or space to the pixel to be predicted.

It should be noted that early in the development of MPEG2, an invitation was made to all interested parties to submit proposed coding algorithms for consideration. Some 32 algorithms were submitted for evaluation. They included subband coding, wavelets, and DCT. Each proponent simulated his algorithm on standard video sequences. The resulting decoded sequences were rigorously evaluated by juries. The result of this evaluation was that the MPEG1 algorithm performed about as good as any other submitted. This was indeed a fortunate result for the cause of compatibility.

MPEG2 is currently under development. The schedule calls for the technical parts of the standard to be frozen in March 1993, with several years additional required to resolve editorial issues and go through the approval process.

The importance of MPEG2 to HDTV is that some of the proponents seem to be trying to bring their systems into alignment with MPEG2, either by modifying their systems to agree with MPEG2, or by trying to get MPEG2 to adopt their proprietary algorithms. An example of the latter is AT&T's proposing that the Vector Quantization approach that they use in their HDTV approach also be adopted by MPEG2.

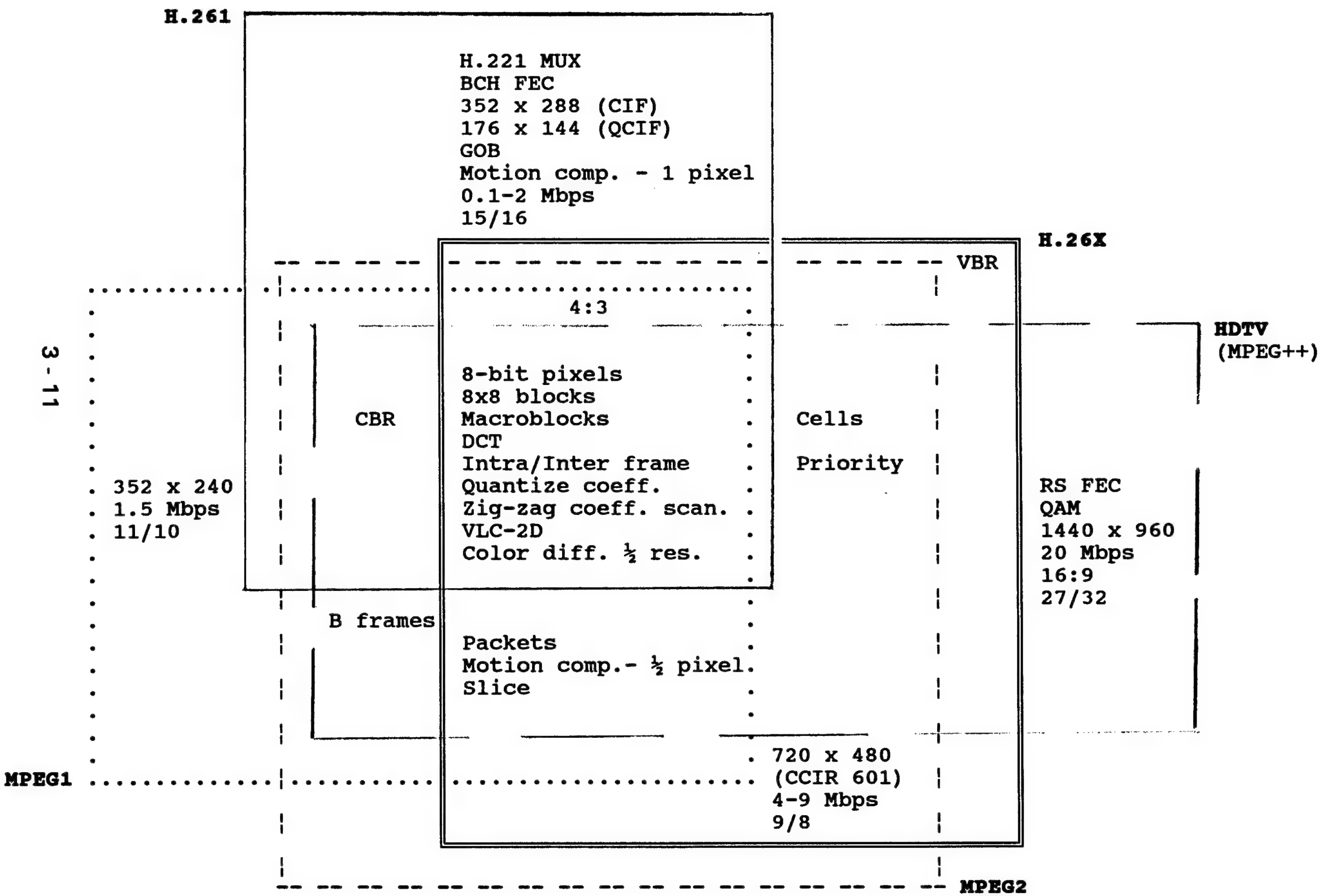
3.3 COMPARISON OF DIGITAL TV STANDARDS

Figure 3.6 is a Venn diagram that shows the elements of the various standards, and indicates which of them are common to two or more. The proposed ADTV (MPEG++) is used to represent HDTV, since it has more in common with the standards than any of the others. It should be noted that this is a rough indication only, since many of the parameters are variable, but are shown in their nominal settings. For example, all but H.26X are shown as CBR, even

though they can and do operate with changing rates, as can occur when other data, such as audio, is switched on or off. And of course H.26X can operate at CBR, if that is appropriate. Also caution should be used for MPEG2, since it has not yet been finalized.

From Figure 3.6, it can be seen that the major differences in the standards are at high levels, such as modulation, multiplexing, FEC, and picture format. The basic low-level video processing is remarkably constant from one standard to another. The consequence of this is that processing "engines" can be built in silicon that can work with any of the standards, and even with others such as JPEG. This will result in great economies of scale and competition that will make feasible economical implementations of all the standards. Even more important, it means that equipment can be built that can work with all the standards, including HDTV, by relatively easy software modifications at the higher levels.

Figure 3.6
Venn Diagram Illustrating Common Elements
of Digital Video Coding Standards



KEY

BCH	Bose-Chaudhuri-Hocquenghem
CBR	Constant Bit Rate
CCIR 601	Standard for Digital TV
CIF	Common Intermediate Format
DCT	Discrete Cosine Transform
FEC	Forward Error Correction
GOB	Group of Blocks
Mbps	Mega bits per second
MUX	Multiplex
QAM	Quadrature Amplitude Modulation
QCIF	Quarter CIF
RS	Reed-Solomon
VBR	Variable Bit Rate
VLC	Variable Length Coding
n x m	Luminance format (pixels x lines)
p:q	Picture aspect ratio (width:height)
r/s	Pixel aspect ratio (height/width)

4.0 HDTV APPLICATIONS IN THE FEDERAL GOVERNMENT

It is generally acknowledged that the use of television throughout the Federal government is pervasive and most critical to its successful operation. When the domestic TV standard was upgraded from monochrome to color, there was a major breakthrough for the Federal government. It is anticipated that the conversion from NTSC to HDTV will be at least as revolutionary. The higher resolution will provide a far more effective link between Federal employees and the vast quantities of information which is generated throughout the Federal government on an ever-expanding scale. Key application areas include multimedia desktop workstations, training/distance learning, simulation, command/control, reconnaissance/surveillance, and video teleconferencing. These applications are discussed below.

Multimedia Desktop Workstations

With the advent of HDTV, the cost of high resolution computer displays on the desktop will drop radically. This will permit the cost effective merging of several media (facsimile, videophone, teleconferencing, computer I/O) at the desktop. Multimedia services involve the integrated delivery and control of multiple information types within a single service, e.g., video, audio, imagery, and test. Multimedia communication will become increasingly important as multimedia applications become more prominent in workstation and personal communications environments. The basic framework for a desktop multimedia workstation is a display having high resolution comparable to the HDTV formats. There is therefore a strong linkage between the multimedia workstation environment and HDTV.

Many of the functions previously associated with the multiplexing of components of a multimedia connection, and embodied in end-to-end signalling and framing systems, such as Recommendation H.221, are provided by basic ATM functionality. Therefore the cost-effectivity of multimedia desktop workstations will increase sharply with the advent of HDTV and broadband ISDN.

Training & Distance Learning

Television plays a major role in the training of all Federal employees, but is

particularly important in the DoD. It is claimed by some that the Desert Storm success was due largely to the excellent training throughout the DoD.

In many cases, the quality of the training mission is impaired by the low resolution of NTSC. For example, in many cases, it is desired to present detailed graphics and diagrams to observers. HDTV will significantly enhance this capability. In the training application, the development of HDTV storage/retrieval equipment is as important as the cameras and displays.

Using the HDTV aspect ratio of 16:9, it becomes easy to include subtitles, data, and other information as explanatory material on either the right or left side of the screen, or to divide the picture into two halves and display them side by side.

In general, the quality of the training increases as the student has the impression of being "present" with the teacher. HDTV creates a greater sense of presence than NTSC.

Simulation

Simulation activity has been rapidly expanding in the DoD in recent years because, for example, it is far less expensive to "simulate" cockpit imagery than to fly airplanes. But the quality of the simulators has been restricted by the resolution of TV displays having reasonable cost. HDTV will radically improve the performance and cost-efficiency of simulation work in the DoD and other Federal agencies.

Command and Control

Television plays a critical role in the performance of the command and control function throughout the DoD. The use of TV for battlefield status displays is absolutely critical, and there is no question that HDTV will provide a quantum leap in the ability to access, process, and most importantly, display a large quantity of command/control information in an effective, practical and inexpensive manner. HDTV will impact the DoD command/control function at the desktop as well as large command centers.

A typical command and control function involves the transmission and display (desk monitor or large display board) of a map with superimposed text and

vector lines. The HDTV format is obviously ideal for this application. It is useful to note that in many situations only a small part of the image changes at one time. In such a case, the DCT/Motion Compensation coding technique could readily transmit the HDTV signal using existing telecommunications or ISDN channels.

Assume an HDTV frame requires 10^6 bits for definition (0.5 bits/pixel). If the channel rate were 384 Kbps, an image can be completely updated in approximately 3 seconds. If only a small part of the image changes, that change can be accomplished virtually instantaneously.

Reconnaissance/Surveillance

The DoD is capable of acquiring vast quantities of reconnaissance and surveillance information from satellites, aircraft, and ground based platforms. In many cases, where it is difficult to store and display this information cost effectively to commanders and analysts, HDTV will significantly help address this problem.

Video Teleconferencing

The purpose of video teleconferencing (VTC) is to provide a sense of visual presence between groups of people in remote locations. One of the disadvantages of existing VTC systems is that this sense of presence is limited in two ways. First, if there is a large group of people in a conference room, the image of one person usually occupies a small fraction of the screen and therefore is not seen clearly. The use of HDTV in such a situation would greatly improve the ability to sense the facial expressions and body language of persons in the remote location.

The second problem with low TV resolution in existing systems is the inability to reproduce high resolution graphic scenes with the required clarity. At the present time, an 8½ x 11 document is typically reproduced by the video camera in two steps -- top half first, then the bottom half. An HDTV system would permit the entire document to be viewed on one screen which would be an important improvement.

5.0 CONCLUSIONS

The following conclusions are drawn from the work performed on this project.

- There is a general trend toward the adoption of a domestic standard for HDTV transmission based upon all-digital technology. At the present time, there are four proponents of all digital systems as listed below.

PROPONENT TEAM	SYSTEM	CHROMA PIXELS
AT&T, ZENITH	SPECTRUM COMPATIBLE	360 X 640
GENERAL INSTRUMENT, MIT	DIGICIPHER	480 X 352
SARNOFF, NBC, PHILIPS, THOMSON	ADVANCED COMPATIBLE TV	480 X 720
GENERAL INSTRUMENT, MIT	ATV PROGRESSIVE	360 X 640

  All four proposed systems employ DCT coding (8 x 8 pixels) and motion compensation, which is similar to the coding technique employed in CCITT Recommendations H.261, H.26x, MPEG-1, and MPEG-2. This trend toward all-digital transmission and alignment with existing standards is obviously a very favorable development for the federal community interested in telecommunications.

- There are two standardization activities proceeding in parallel which are highly inter-related; (1) the U.S. HDTV standard developed by the FCC, (2) the international MPEG-2 standard by ISO. The two organizations developing these standards are attempting to achieve as much compatibility as possible. It is planned to freeze the specifications for both standards in 1993.

- On the international front, the Japanese are providing the MUSE system (1125 lines; 60 fields/sec.) on an operational tariffed basis,

while the Europeans are experimenting with the MAC system (1250 lines; 50 fields/sec.). While these systems are both fundamentally analog in nature, there is considerable concern in these foreign countries that the U.S. all-digital standard will supersede their approaches. For this reason, they are supporting work by the CCIR to develop an international all-digital HDTV standard.

- It is anticipated that the HDTV system will have a major impact within the Government community. Examples of important applications are multimedia desktop workstations, training/distance learning, simulation, command/control, reconnaissance/surveillance, and video teleconferencing.

6.0 RECOMMENDATIONS

Based upon the work performed on this project, the following recommendations are proposed.

- To maximize interoperability between future HDTV systems and Federal telecommunication systems (e.g., video teleconferencing), it is recommended that the domestic HDTV standardization process and the international MPEG-2 standardization efforts be aligned as closely as possible. To accomplish this alignment, it is recommended that Delta continue its active participation in both of these organizations. To insure that the H.26x Broadband ISDN video coding standard is closely aligned with these standards (HDTV, MPEG-2), it is further recommended that Delta continue its activity with the ATM Video Coding Experts Group. For example, it is important that the MPEG-2 standard have the potentiality for low delay and high error resilience to permit an easy extension to H.26x.

- To maximize the interoperability and applicability of the future HDTV system throughout the Federal government, it is recommended that the selected system be compatible with progressive scan and that packetized transmission be employed. This will maximize compatibility with computerized systems and future ATM networks.

APPENDIX A

INTEROPERABILITY ASSESSMENTS

Reference 1

INTEROPERABILITY ASSESSMENTS

December 11, 1992

StellaCom, Inc.
Advanced Television Systems
1901 N. Moore Street - Suite 601
Rosslyn, VA 22209-1706

(703) 243-1695 (p)
(703) 243-1698 (f)

This work was done for ACATS PS/WP-4.

INTRODUCTION

This document was written by StellaCom, Inc. for PS-WP/4, a working party of the Advisory Committee on Advanced Television Systems (ACATS) which operates on behalf of the Federal Communications Commission. The information contained in this document was obtained from documents submitted by the proponents to PS-WP/4 and to other working parties of ACATS, and from oral presentations by the proponents in September, 1992 in which questions from PS-WP/4 were answered.

It was the intention of PS-WP/4 for this document to specify whether each feature discussed was

1) in the system tested at ATTC,

2) promised for the system to be field tested,

3) promised for the system after field testing, or

4) possible/realistic at a future time.

This was done to the extent possible with the information available. However, the proponents are continuing to enhance their systems and revise their time estimates for the implementation of features. Some of the enhancements under way now are in response to the concerns of PS-WP/4 on Interoperability, Scope of Services and Features, and Extensibility.

Financial support was provided by Apple Computer, Inc., Bellcore, Cable Television Laboratories, Inc., Eastman Kodak Company, Digital Equipment Corporation, and Viacom International, Inc.

Advanced Digital HDTV Interoperability Assessment

System Description

The Advanced Digital HDTV (AD-HDTV) System from the Advanced Television Research Consortium is an all-digital system with 1050 scan lines per frame and 29.97 frames per second. Compression and transmission are in a 29.97 hz progressive form. The system tested used a 2:1 interlaced source and a 2:1 interlaced display. The display format of the prototype system has a 16:9 image aspect ratio with 960 lines per frame and 1248 pixels per line, although future versions are expected to use 1440 pixels per line. The proponent also claims that the system can be changed to have square pixels. Features of the system include a prioritized data transport format that separately packages video, audio and data and allows their mix to vary dynamically. It also includes two-tier prioritization where each data segment is assigned to a high-priority (HP) data stream or a standard priority (SP) data stream. Transmission is by Spectrally Shaped QAM (SS-QAM) consisting of two 4- or 5-bit-per-symbol QAM signals fitted between co-channel NTSC carriers. Each data stream modulates a 16- or 32-QAM carrier, with the high-priority carrier transmitted at higher power. This allows a degree of graceful degradation on the edge of the broadcast contour and transmission robustness within the service area, while minimizing interference to and from co-channel NTSC. It also allows a degree of bit-stream scalability.

I. Interoperability

A. <u>with Cable TV</u>

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The question remains whether HDTV receivers can be built to support higher numbers of bits per symbol without cost impact.

Encryption and addressing are important service features for Cable TV. See Section II, items F and G.

B. <u>with digital technology</u>

Since this system is all-digital, the advantages of all-digital systems apply.

C. <u>with headers/descriptors</u>

The AD-HDTV system formats all its data in a layered manner following the MPEG syntax. After synchronization, the data link layer identifies the service type. Within the adaptation level, the adaptation header contains information governing the packing of variable-length code words for video, and information used in recovery after channel changes or errors. The video service level includes the actual encoded video information.

All data sent by the AD-HDTV system is grouped into fixed-length cells that contain data of a single particular type. The cells are 148 bytes long, starting with a single byte for synchronization, an 8-bit Barker sequence {00011101}. This is followed by a single byte for a service header that defines cell

types at the link layer. Next is a 4-byte adaptation level header that contains pointers to the start of a slice within the picture frame, as well as the identity of the slice within the frame. These are followed by a 120-byte information payload containing actual data for the cell's particular service type. At the end of the cell is a 2-byte frame check sequence (FCS) trailer, and a 20-byte forward-error-correction (FEC) trailer.

The adaptation layer pointers and slice identification information provide re-entry points within the codec video data, making it possible to begin decoding at a known point after an error event or channel change that requires a restart for some or all of the video decoding.

D. with NTSC

The proponent selected the field rate of 59.94 Hz for compatibility with NTSC. The number of active video lines, 960, was selected to be double the number of active NTSC lines, 480. (NTSC normally has 483 active video lines, but 483 is an awkward number for conversions.) Since the aspect ratio of this system, 16:9, is different from the aspect ratio of NTSC, 4:3, a choice must be made for the conversion. Two choices for down-conversion are 1) "Edge Crop", in which the HDTV picture fills 480 lines on NTSC with loss of the sides of the HDTV picture, and 2) "Letterbox", in which the full width of the HDTV picture is displayed in the full width of the NTSC picture, but leaving unused areas at the top and the bottom of the NTSC screen. Pixel values from HDTV lines are stored in memory and read out at reduced speed to make NTSC lines. The range of pixels read and the clocking rate depend upon the choice of edge-crop or letterbox. The down-conversion involves interpolation between HDTV pixels in a line and between HDTV lines.

E. with film

It is claimed that AD-HDTV will support an "electronic film" format that makes use of the redundant field to improve quality. Since film has a slower (24 frames per second) temporal rate, AD-HDTV will scan its 1050-line raster progressively for film at 24 frames per second, with the same 1440 x 960 pixel format that it uses with video sources. However, in film productions where computer graphics are used extensively, square pixels may be more desirable in the image representation. It is claimed that AD-HDTV will also provide a progressively scanned 1440 x 810 square pixel format to accommodate this source. Most receivers will perform 3:2 pull-down to convert to their 59.92 Hz field rate, but higher-cost receivers could use 3:1 frame repeat to display at 72 Hz.

F. with computers

Progressive scanning and square pixels are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations.

Encoding and transmission in AD-HDTV are done in progressive form, favoring interoperability with computers, although testing of this system has been done with interlaced sources and displays requiring format conversions. ATRC has suggested that the system will eventually use progressive sources and displays. With 1440 pixels/line, the pixels in AD-HDTV are 18.5% wider than high, and

with 1248 pixels/line they are 37% wider than high. ATRC has suggested that their system can provide square pixels by reduction of the number of active scan lines to 810.

This system also offers the possibility of interoperability with MPEG-based computer applications. See Section J.3 below.

G. with satellites

Digital signals with bit rates used by the proponents for terrestrial transmission can generally be accommodated in the available bandwidth of a satellite transponder. The total data rate for the AD-HDTV system in the mode tested by ATTC is 24.0 Mbps. The proponent has suggested removing the 0.9-rate trellis code used with SS-QAM, making the net data rate 21.6 Mbps. The proponent does not anticipate the need for any additional error correction for satellite transmission, although convolutional coding is normally used. The proponent stated that three AD-HDTV channels may be carried in a transponder. However, it is unlikely that more than two AD-HDTV channels will be carried in a typical 36-MHz transponder. See the link analysis done for the DSC-HDTV signal. The proponents also point out that it is possible to carry AD-HDTV and NTSC signals on the same transponder.

H. with packet networks

The data link sublayer format is based on a "cell relay" asynchronous time-division multiplexing concept similar to the asynchronous transfer mode (ATM) standard that will be used for broadband integrated services digital network (B-ISDN). The data-link header contains information such as priority indicator, service ID and cell sequence number. This provides service-independent transport services such as priority support, service multiplexing, and cell-error detection and correction. For the received bit-stream, the transport decoder performs Reed-Solomon decoding and a cyclic redundancy check (CRC) for error detection. Cells received in error after correction are discarded by the demultiplexer. Each cell contains 148 bytes.

The proponent claims that minor picture artifacts may be noticed at a packet error rate of 10^{-3}. However, AD-HDTV will make viewable pictures and perfect sound even if only its High Priority cells are received. This is effectively an error rate of 8 out of 10 packets lost. Out-of-order packets can be put back in order at the receiver.

I. with interactive systems

The total rate control buffer delay is nominally 250 msec for both the high priority and standard priority buffers. In addition, there may be image build-up after a switch or a scene cut depending on the place in the 9-frame Group-of-Pictures (GOP) cycle. Total acquisition time has been designed to be limited to a worst case value of about 1 sec. This number is based on 0.5 sec for the modem and equalizer and 0.5 sec for the video/transport acquisition.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. According to the proponent, a full-quality AD-HDTV encoder requires 4 frames of latency, due to its use of bi-directional motion compensation with an MPEG GOP structure of M=3, N=9. An additional frame is needed for interlaced-to-progressive conversion. Similar delays are present at the receiver. Rate buffers at the transmitter and at the receiver cause additional delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

This proponent claims that for interactive applications where latency is a concern, an encoder can produce an MPEG bit stream using only forward motion compensation to reduce the coding part of the latency.

J. Format conversion

1. with 1125/60

Up-converting to the Common Image Format (1920 x 1080) requires a 8:9 vertical interpolation and a 3:4 horizontal interpolation from 1440 x 960. SMPTE 240M currently uses 1035 active lines. Colorimetry used by AD-HDTV is intended to be consistent with SMPTE 240M.

2. with 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Simple methods that involve frame dropping lead to jerky motion, but other techniques produce acceptable images under most conditions. This difficult conversion is not simplified by the fact that both the source system and the target system are interlaced 2:1.

3. with MPEG

AD-HDTV's use of MPEG-based video and audio compression provides the possibility of interoperability with MPEG-based computer multimedia applications directly in the compressed bit-stream format. The underlying video compression algorithm adheres to the MPEG1 standard from ISO, in that parameters allowable within the MPEG1 definition are used although they are not the MPEG1 default parameters. Prior to entering the prioritization and transport processors, the compressed video representation produced at the encoder fully conforms to the MPEG1 specification. An MPEG bit-stream can be obtained from the output of the compression encoder at the interface to the priority processor. Note that the AD-HDTV encoder prototype uses an internal fixed-length representation for MPEG code words at the interface between its compression and prioritization stages, so that a standard MPEG1 bit stream is not currently available as an output from the prototype hardware.

4. with still image

AD-HDTV's MPEG1 compression, based on the DCT, is generally compatible with JPEG, the ISO standard for still-image compression.

Photo CD uses a proprietary subband coding approach for compression. Picture interchange with AD-HDTV requires full decompression and recompression.

CDI uses the MPEG1 compression syntax for motion video. Therefore, AD-HDTV will be directly compatible with CDI.

K. Scalability

Easy scalability is claimed by the proponent. The MPEG1 bit-stream produced by the encoder is organized into several layers. The ISO-MPEG standard not only specifies the form of the compressed video bit-stream, but it incorporates several higher layers for specifying video parameters. In particular, the "sequence header" provides a mechanism for dynamic specification of the picture size (in terms of pixels x lines), the pixel aspect ratio and the frame rate. With this mechanism, AD-HDTV supports dynamic reconfiguration among its nominal and optional spatial resolutions and raster formats (e.g., interlaced HDTV and progressively scanned film modes). These capabilities are not implemented in the AD-HDTV prototype hardware tested at the ATTC.

The picture produced by AD-HDTV's HP signal alone is a reduced-quality image. The proponent has announced his intention to reprioritize the data partition to make the HP-only picture useful as a

digital hierarchy. The present form is considered by the proponent as sufficient to carry the viewer through short SP channel outages, and as a base for error concealment for high SP error rates. The decoded artifacts observed in an HP-only reconstruction will depend on the exact priority processing algorithm. Typical priority processor operation results in a lower resolution and slightly jerky motion rendition since the B-frame data is carried primarily in the SP channel.

For picture-in-picture and picture-out-of-picture, only the reprioritized HP signal with its 4.8 Mbps data steam needs to be processed, with savings in hardware over full AD-HDTV processing.

For multiple programs in a single channel, AD-HDTV's Prioritized Data Transport layer provides for asynchronous delivery of multiple service types. Multiple video streams can be assigned individual service types.

II. Scope of services and features

A. Initial Use of Ancillary Data

Because of AD-HDTV's asynchronous data multiplexing, there is no hard partitioning of ancillary data needed. Unassigned service types provide for the delivery of many types of ancillary data.

B. Audio

AD-HDTV offers flexibility in the mix of audio, video and data and therefore the number of audio channels. The audio system chosen for AD-HDTV uses a coding algorithm called MUSICAM, which is related to standards endorsed by the ISO in conjunction with MPEG. The audio coding accepts 16-bit samples of audio at a rate of 48 k-samples/second, for each channel of the stereo pair. The MUSICAM algorithm compresses the 1.54 Mbps to 256 kbps. AD-HDTV provides an additional 5 dB of robustness for audio, as it transmits audio as High Priority data. Thus, audio remains when only the HP picture can be decoded.

C. Data

An allocation of 256 kbps has been made for auxiliary data, and that data has been set aside within the prototype equipment as SP data. A standard communications interface port for this auxiliary data was provided for the prototype.

Auxiliary data is normally sent in standard priority (SP) mode. The actual priority used in an operational system is a broadcaster option and will depend on the type of data service. If <100 kbps of HP capacity were used for data, the proponent claims that the impact on HP-only picture quality would be small; for higher auxiliary data rates, use of SP is recommended, since additional reliability in data services can be provided via appropriate retransmission schedules, etc. Also, see G. Addressing.

D. Text

Text delivery is discussed below in Section I on Overlay.

E. Captioning

It was suggested that 9.6 kbps of the ancillary data capacity be used for closed captioning.

F. Encryption

The prototype hardware did not include encryption, but AD-HDTV's transport structure has been designed to readily accommodate encryption. The proponent expects to detail the encryption method with industry participation.

G. Addressing

The system provides opportunities for high-data-rate burst-mode delivery of auxiliary data. As a limit, the entire channel capacity, 18.5 Mbps, could be dedicated to addressing a large number of receivers with decryption keys during the few seconds preceding a popular pay-per-view program. A still picture with audio using 4 Mbps could be transmitted while 14.5 Mbps is devoted to subscriber addressing. Conditional access data can be treated as a special service type and packaged into its own transport cells, or included within the video and audio data. It can be decoded without decoding video and audio.

H. Low-cost receiver

AD-HDTV is partitioned into High Priority and Standard Priority data streams that modulate separate carriers. A low-cost receiver can be built to use only the HP stream with savings on the tuner and processing components. The HP-only mode has been demonstrated, showing high resolution but poor response to motion. The proponent anticipates improvements to the prioritization algorithms to make HP-only pictures that are much better suited for small, low-cost receivers.

I. VCR capability

No hardware development of VCR's has been reported. However, it is thought that digital recording is within current technology for consumer use.

MPEG has periodically occurring frames that are entirely spatially coded. This is said to provide the ability to reconstruct pictures in fast-forward and reverse scanning modes from digital storage media.

MPEG compression has periodic spatially-coded frames, allowing splices and inserts to be made on GOP boundaries in the compressed transmission format. Limited picture cropping can be handled in compressed form if it aligns with macroblock boundaries.

AD-HDTV has provision for carrying text and graphics overlay data. This data can be sent as a separate service type, to be superimposed on the display at the receiver. The proponent recommends that receivers have, as a minimum, overlay capability for simple text. This would permit the local broadcaster to insert text of local importance into the data stream containing the compressed signal that he passes through from the network source to the viewers, without decompressing and recompressing the video.

III. Extensibility

A. to no visible artifacts

AD-HDTV has, as its core, the MPEG1 compression standard that is extensible to virtually any data rate. The proponent points out that AD-HDTV at 17.7 Mbps is already an extension of the baseline MPEG1 parameters which encode low-resolution video at 1.5 Mbps.

B. to studio-quality data rate

AD-HDTV was designed with the anticipation of several levels of related MPEG compression. The proponent suggests that a studio standard could be set at 216 Mbps, the data rate of existing studio D-1 recorders. He believes that this standard would be of extremely high quality and would permit studio processing without degradation.

C. to higher resolution

The proponent claims that AD-HDTV potentially supports the delivery of other video and image formats over appropriate bandwidth channels to special receivers with increased memory. The MPEG1 core allows resolutions up to 4095 x 4095. Examples given by the proponent follow. One is a 1280 x 1024 progressive scan, logically square pixel format that is appropriate for computer workstations. Another is a 65,520 x 65,520 very-high-resolution still-image format (the limit of JPEG addressing capability). Even higher resolutions can be delivered by extending the JPEG addressing capability. Such formats would likely prove useful in medical imaging, computer graphics, and space and defense applications.

The proponent has discussed the possibility of introducing ultra-high-definition television by sending augmentation data packets assigned to a unique service type that will be disregarded by older receivers but processed by new receivers.

D. Provision for future compression enhancement

The proponent suggests improved calculation of motion vectors and improvements in bit allocation and prioritization as likely means for picture quality improvement without changing the receivers or the data rate.

Channel Compatible DigiCipher HDTV Interoperability Assessment

System Description

The Channel Compatible DigiCipher (CCDC) HDTV System, also known as the Advanced Television Alliance (ATVA) Progressive System, is an all-digital system that is progressively scanned with 787.5 lines per frame and 59.94 frames per second. The display format has square pixels with 720 active lines by 1280 pixels per line in a 16:9 image aspect ratio. Several types of digital coding are used for compression and signal robustness. Transmission is by quadrature amplitude modulation with 4 or 5 bits per symbol, offering the broadcaster the choice of two data rates involving a tradeoff between picture quality and coverage area. The system accommodates either 16- or 32-QAM modulation. The proponent emphasizes that the system was designed to be highly modular, so that the video processing, the audio processing, and the transmission system can be used independently.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The question remains whether HDTV receivers can be built to support higher numbers of bits per symbol without cost impact.

Encryption and addressing are important service features for Cable TV. See Section II, items F and G.

B. with digital technology

Since this system is all-digital, the advantages of all-digital systems apply.

C. Headers/descriptors

A frame header identifies the video source material, the frame rate, resolution, aspect ratio, and other system data.

D. with NTSC

As the CCDC system is directly related to NTSC, transcoding to NTSC is straightforward. Edge-crop conversion to CCIR 601 involves discarding 160 HD samples from each end of the horizontal line and 4:3 interpolation of the remaining samples. It can involve 3:1 and 2:1 vertical interpolations in alternate fields. For letter-box, a 16:9 horizontal interpolation is required, as well as a 3:1 vertical interpolation of the active lines. For edge-crop display, the 480 lines can be placed in a specific set of the 483 active lines. The first (odd) field can be displayed on lines 21, 23, ..., 499, and the second (even) field is displayed on lines 22, 24, ..., 500. The active lines 501 and 503 of the odd fields and 502 of the even fields can be displayed as black. For letter-box display, the vertical interpolation is more

complicated. Up-conversion from NTSC requires line tripling, horizontal line-rate conversion and interpolation.

E. with film

The prototype has the following film mode. Film is displayed with the three-two pulldown process for 24 frame/sec film and with simple frame repetition for 30 frame/sec film. The proponent claims actual frame rates of 59.94, 30 and 24 frames/second, implying that no film slowdown is needed to accommodate the standard video frame rate. The encoder automatically detects the presence of 24 frame/sec or 30 frames/sec scene material from film sources. When a film source is detected, an alternate buffer control algorithm is used which takes advantage of repeated frames in the source. With the scanning method used in CCDC, only two out of each five TV frames need to be sent for 24 frame/sec film. The proponent claims that the video quality in the film mode is essentially equivalent to the uncoded source. He also claims that full-resolution color is possible in the film mode, though that capability is not in the prototype.

F. with computers

Progressive scanning and square pixels, both used in this system, are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production, and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Also, scan conversion between interlaced and progressive systems can produce undesirable artifacts. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations.

G. with satellites

Digital signals with bit rates used by the proponents for terrestrial transmission can generally be accommodated in the available bandwidth of a satellite transponder. The total data rate for the CCDC system in the mode tested by ATTC is 26.4 Mbps. The proponent suggests that 8-PSK modulation would permit two CCDC signals per 36 MHz transponder. However, normal transmission by satellite is QPSK (4-phase). CCDC, which has the highest data rate of the all-digital proponents, would be the most difficult to transmit with two channels per transponder. Nevertheless, using the 19.9- Mbps information rate of CCDC, Reed-Solomon coding, and rate 7/8 convolutional coding, two channels can probably be transmitted in a 36-MHz transponder.

H. with packet networks

CCDC data is organized into 525 data lines per frame. These data lines could be used as packets if augmented with packet assembly information. Error concealment, already implemented, would ensure some resistance to packet loss. The proponent has suggested implementation of a packetized structure after field testing.

I. with interactive systems

A form of interactivity important to cable operators is the speed with which a customer can scan through many channels. This system is reported to acquire a channel in 0.4 seconds, providing the sound and beginning to build the picture. The picture is at full resolution in 0.73 seconds.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. Frame delays are required at the transmitter and at the receivers for coding and decoding. Rate buffers at the transmitter and at the receiver cause additional delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

CCDC reports a video delay of 5 or 6 frames, corresponding to 83 to 100 ms. The exact time is said to depend on how the frame buffer is used, with the video/film selection a factor.

J. Format conversion

1. 1125/60

Upconverting to the Common Image Format (1920 x 1080) is easily done by 2:3 interpolation horizontally and vertically. SMPTE 240M currently uses 1035 active lines. Colorimetry is SMPTE 240M.

2. 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Present methods that involve frame dropping lead to jerky motion, but other techniques can produce acceptable images under most conditions. This difficult conversion may be easier from a progressive source than from an interlaced source.

3. MPEG

There is no direct compatibility in terms of bit stream. However, see the MPEG discussion for DigiCipher.

4. with still image

The capture of still images from video is favored by progressive scan.

K. Scalability

The proponent claims that the system is scalable in the sense that the decoded signal may be interpolated or decimated to suit various displays. Scalability by picture interpolation can be implemented in any proposed system. It is simplified by the progressive scanning in this system. The system may also be scalable in the sense that a subset of the bits, such as those that correspond to the lower-order discrete cosine coefficients, could be used by lower-cost VLSI to obtain a lower-resolution signal. However, the entire bit stream must be available for the subset to be extracted, and the other digital systems have similar capabilities. Picture-in-picture and picture-out-of-picture are handled by standard methods in the receiver.

II. Scope of services and features

A. Initial Use of Ancillary Data

Assignments have not yet been made. See Data.

B. Audio

The CCDC system is designed to have four channels of audio using 503 kbps or six channels of audio using 755 kbps. Unused audio capacity is used to carry additional video data. The audio system uses a coding algorithm called MIT Audio Coder, presently implemented with Motorola DSP-96002 floating-point digital signal processors. It accepts 16-bit audio samples at a rate of 48 kHz for each channel. All audio channels are coded independently. Capacity has been reserved for dynamic bit allocation for different parameter sets. The proponent claims that this will permit the system to incorporate future improvements in audio compression without changing the receivers.

C. Data

Auxiliary and control data has been allocated 251.8 kbps. This may be implemented with an RS-422 interface. In the current hardware, the only access to the auxiliary data channel is via four asynchronous 9600-bps RS-232 interfaces. However, the design can be modified to permit data capacity, not used by audio, to be reassigned to data services.

D. Text
E. Captioning

Teletext and captioning is sent in the ancillary data channel. Additional ancillary data capacity will be provided if needed.

F. Encryption
G. Addressing

The first byte in each data line is reserved for control information, described as including decryption keys and subscriber data. There are 525 data lines per frame and 59.94 frames per second. Thus, there are about 252 kbps of capacity for this kind of data. Encryption has not yet been implemented.

H. Low-cost receiver

This system will permit lower receiver cost using lower-resolution displays. The proponent claims that progressive scan used in this system favors such scalability. The low-cost receiver would have to extract the information it can use from the complete data stream, perhaps just the DC DCT coefficients, but would save on processing capability as well as display cost. The proponent claims that the <u>standard</u> HDTV receiver can be built with 12 custom VLSI's and 2 Mbytes of memory.

I. VCR capability

The proponent reports no hardware development of VCR's specific to CCDC, but refers to the DigiCipher/Toshiba VCR that has been demonstrated within ATVA. The CCDC data stream, about 20 Mbps, is within the capability of current technology for consumer use.

It is claimed that a rapid search mode can be implemented by reconstructing the images from those blocks coded with no temporal predictor. This gives at least three displayable frames for every

60. Additional intra-coded blocks may also be used as they occur. The resultant picture would have full resolution, but may include artifacts. The reverse playback cannot be done with full quality because predicted frames cannot be generated.

Splice and insert could be handled by forcing the receiver to reacquire. Crop and overlay would require that the data stream be decompressed first.

Square pixels and progressive scanning simplify the implementation of special effects such as zooming and panning.

III. Extensibility

A. to no visible artifacts

Based on simulation tests, the proponent believes that the compression algorithm will produce no visible artifacts at a data rate of 50 Mbps, regardless of the difficulty of the camera-generated source material.

B. to studio-quality data rate

According to the proponent, the intraframe encoding mode for the whole frame can be used for a production standard. Here, every frame is encoded without motion prediction. Production-quality video with a resolution of 1280 x 720 can be stored with 3 Mb/frame using the intraframe compression method included in this system. At 60 frames per second, the bit rate is 180 Mbps, an acceptable rate for studio use. The proponent claims that the frame can be decoded and re-encoded many times with little degradation.

C. to higher resolution

Currently the system is designed to display 1280 x 720 image sequences, but larger sizes can be specified as part of the frame header.

D. Provision for future compression enhancement

The compression algorithm can be improved by performing better motion estimation, and including better perceptual criteria at the transmitter. These involve no changes at the receiver.

DigiCipher HDTV Interoperability Assessment

System Description

Also known as the Advanced Television Alliance (ATVA) Interlaced system, the DigiCipher HDTV system is an all-digital system with 1050 scan lines per frame, 59.94 fields per second, and 2:1 interlace. The display format has a 16:9 image aspect ratio with 960 lines per frame and 1408 pixels/line. Several types of digital coding are used for compression and signal robustness. The primary picture coding is discrete cosine transform with motion compensation. Transmission is by quadrature amplitude modulation with 4 or 5 bits per symbol, offering the broadcaster the choice of two data rates involving a tradeoff between picture quality and coverage area. The system accommodates either 16- or 32-QAM modulation.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The proponent states that it is likely that a rate higher than 32-QAM will be used for multichannel NTSC, and has expressed the intent to maintain compatibility between DigiCipher Multichannel NTSC and HDTV. The question remains whether HDTV receivers can be built to support higher numbers of bits per symbol without cost impact. The proponent has conducted demonstrations of one-way and two-way operation over existing cable TV plants.

Encryption and addressing are also important service features for Cable TV. See Section II, items F and G.

B. with digital technology

Since this system is all-digital, the advantages of all-digital systems apply. The proponent has agreed to determine the feasibility of a common standard between the DigiCipher digital bus and the EIA CE bus.

C. Headers/descriptors

The proponent discussed the use of the ancillary data space for transmitting the program name, remaining times and program rating. In the prototype, there is a 7-byte header at the beginning of each data frame. Four bytes are used, and three bytes are available. There is a one-byte header at the beginning of each video frame, of which one bit is available. There is a fully defined two-byte header at the beginning of each macroblock. In his certification disclosure, the proponent discussed his intention to implement adaptive allocation between video, audio and ancillary data. The proponent has now announced an improvement involving expansion of the video frame and macroblock headers, and packetization at the transwitch layer. Transwitch layer packetization makes the data allocations adaptive, allowing flexible allocation of capacity among video, audio, and data applications. The proponent states that this modified data format will be implemented in the prototype prior to field testing.

D. with NTSC

The proponent selected the field rate of 59.94 Hz for compatibility with NTSC. The number of active video lines, 960, was selected to be double the number of active NTSC lines, 480. (NTSC normally has 483 active video lines, but 483 is an awkward number for conversions.) Since the aspect ratio of this system, 16:9, is different from the aspect ratio of NTSC, 4:3, a choice must be made for the conversion. The limiting choices for down-conversion are 1) "Edge Crop", in which the HDTV picture fills 480 lines on NTSC with loss of the sides of the HDTV picture, and 2) "Letterbox", in which the full width of the HDTV picture is displayed in the full width of the NTSC picture, but leaving unused areas at the top and the bottom of the NTSC screen. Pixel values from HDTV lines are stored in memory and read out at reduced speed to make NTSC lines. The range of pixels read and the clocking rate depend upon the choice of edge-crop or letterbox. The down-conversion involves interpolation between HDTV pixels in a line and between HDTV lines.

E. with film

The prototype system accepts film shot at 24 frames per second as 59.94-Hz video, 2:1 interlaced, having been converted by the three-two pulldown technique. The DigiCipher encoder recognizes the redundancy in each five-field sequence as having originated in 24-frame film, and converts the 59.94-field video back to 23.98 frames/second. The image is processed and transmitted as 23.98 frame progressive, and converted back to 59.94-field interlace in the decoder, using 3:2 pull-down. Future receivers could alternatively use 3:1 frame repeat to display progressive at 72 Hz. Similarly, 30-frame film source which is delivered to the encoder as 59.94-field video is processed and transmitted as 29.97-frame progressive. The benefit is more efficient coding, and thus higher quality. The proponent has proposed a system improvement which would allow receiving and processing images directly in 24- and 30-frame progressive. He also suggests that in the future, spatial resolution could be increased as temporal resolution is decreased.

F. with computers

Progressive scanning and square pixels, not included in the DigiCipher system tested, are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Also, scan conversion between interlaced and progressive systems can produce undesirable artifacts. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations. Non-square pixels do not present a problem for the display of RGB signals from a computer's video card on HDTV receivers, but would complicate more sophisticated attempts at interoperability, such as the display by an HDTV receiver of a picture from a digital data stream generated by a computer. This system has pixels that are 21% wider than high.

The prototype hardware was built to select between field processing and frame processing for each superblock, depending upon its motion, in order to provide optimum motion handling. However, computer interoperability would be enhanced if the encoder were forced to do frame processing on all superblocks. With this feature, coding and transmission would be in progressive form. The proponent has proposed adding this feature as an option at the encoder.

G. with satellites

Satellite transmission of the DigiCipher HDTV signal was recently demonstrated using QPSK in a 24-MHz bandwidth, achieving a raw data rate of 39 Mbps. Instead of the trellis coding used in the terrestrial system, convolutional coding with a Viterbi decoder was used. The coding was rate-1/2, so that the data rate after Viterbi decoding was 19.51 Mbps. Reed-Solomon coding was also used, with the information rate being 18.2 Mbps, identical to the terrestrial signal. A 5.5 dB C/N threshold was achieved; an improvement over the 8+ dB threshold typically achieved in NTSC satellite transmission. The proponent recommends using rate-3/4 coding to yield a 50% increase in the information rate. This would support a higher-level compressed HDTV signal or an NTSC signal sharing the channel with the HDTV signal. In a 36-MHz transponder, two transmission-quality HDTV signals, or alternatively, one distribution-quality, 40-45 Mbps signal can be transmitted.

H. with packet networks

In the prototype tested at ATTC, the video data was packetized, with each packet, about 2 kb long, containing one macroblock of video data. For lost packets, the decoder would have used error concealment which was already implemented to handle transmission errors.

The proponent has since offered a redesign of the data structure with headers/descriptors. The benefits of this approach include improved packet transmission. Packets will be organized by data type. Packet length will be 155 bytes, with the payload being 141 bytes. Out-of-order packets can be reordered in a suitably designed receiver. Implementation is expected after field testing.

I. with interactive systems

A form of Interactivity important to cable operators is the speed with which a customer can scan through many channels. This system is reported to acquire a channel in 0.4 seconds, providing the sound and beginning to build the picture. The picture is at full resolution in 0.77 seconds.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. Frame delays are required at the transmitter and at the receiver for coding and decoding. Additional delay may be needed to facilitate frame coding in an interlaced system. Rate buffers at the transmitter and at the receiver also cause delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

According to the proponent, the latency of DigiCipher is 83 msec.

J. Format conversion

1. with 1125/60

Up-converting to the Common Image Format (1920 x 1080) requires a 8:9 vertical interpolation from 1408 x 960. SMPTE 240M currently uses 1035 active lines. Conversion from 1408 pixels/line to 1920 pixels/line requires an 11-to-15 interpolation. The proponent says that 15 interpolation filters are required, each having 5-6 taps. Colorimetry used by DigiCipher is the same as SMPTE 240M.

2. with 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Present methods that involve frame dropping lead to jerky motion, but other techniques can produce acceptable images under most

conditions. This difficult conversion is not simplified by the fact that both the source system and the target system are interlaced 2:1.

3. with MPEG

The DigiCipher HD decoder would require modification to decode MPEG or H.261. The proponent states that there would be only a modest increase in complexity because DigiCipher shares many commonalties with MPEG and H.261. MPEG and H.261 decoders will not decode DigiCipher HD.

4. with still image

The proponent has identified still-frame as a useful capability, and believes that forward compatibility with JPEG, Photo CD and CDI is feasible and is a marketplace issue. The frame-coding option offered by the proponent (See I.F) enhances compatibility with still images.

K. Scalability

Though the receive and display clocks are currently linked, the proponent proposes to operate them independently in the future. The receiver could then receive non-real-time video at slower rates.

According to the proponent, picture-in-picture and picture-out-of-picture are possible with DigiCipher as receiver design options.

L. Other Compatibility Features

DigiCipher HD processes the image in four parallel panels. Each panel processor is comparable to a DigiCipher NTSC processor and thus is able to process a DigiCipher NTSC signal. There is also a compatible bus that can support both NTSC and HD signals. The proponent claims that the compatibility extends to VCR's, and satellite and cable receivers.

II. Scope of services and features

A. Initial Use of Ancillary Data

The use of the ancillary data capacity is not yet specified.

B. Audio

The DigiCipher system provides for four independent audio channels, each sampled at 47.2 kHz (to be changed to 48 kHz) and digitized to 16-bit precision. The system uses two Dolby AC-2 compression systems operating at 24-bit precision. The two compressed audio data streams are formatted along with a 1200-bps control signal into a single serial output data stream at 503 kbps, to be multiplexed into the transmitted signal. The system description discusses extra error protection within the Dolby AC-2 decoder. This leads to audio that remains acceptable under signal conditions where the picture fails.

The proponent has proposed to incorporate the Dolby AC-3 composite-coded 5.1 channel surround-sound system into the prototype prior to field testing. This system offers a variety of modes of operation, including dual independently coded AC-2A channels. The proponent also suggests that the use of packetization will permit potentially numerous composite-coded and independently coded audio channels to be transmitted, with the receiver determining which to process.

C. Data

A 126-kbps channel is provided for ancillary data. This is distinguished from a separate 126-kbps channel assigned for conditional access use. The planned incorporation of packetization at the transwitch layer will allow the system to flexibly allocate transmission capacity to data.

D. Text

In the prototype, four data channels at 9600 bps were implemented to illustrate asynchronous data transmission.

E. Captioning

The proponent has stated an intention to demonstrate closed-captioning compatibility when prototype captioning hardware becomes available in mid-1993.

F. Encryption

Encryption was not implemented in the prototype. However, the proponent claims to have developed "smart card" security for both VideoCipher and DigiCipher. With packetization by data type, the different types of data will be encrypted separately.

G. Addressing

A 126-kbps channel is assigned for conditional access use (subscriber addressing), separate from the 126-kbps channel for auxiliary data. The proponent suggests that with this data channel, 50-100 million receivers could be addressed in less than one day. The proponent offers to implement fully packetized data that will allow flexible allocation of data with increased peak-load subscriber-addressing capability.

H. Low-cost receiver

The proponent suggests a low-cost receiver that displays 176 x 120 pixels by extracting and processing only the DCT DC coefficients. Picture quality is suggested to be comparable to that of NTSC on a 2" LCD display. Extracting the DCT DC coefficients for a low-cost receiver can be implemented in any of the proposed digital systems.

I. VCR capability

A consumer-grade VCR has been publicly exhibited by GI and Toshiba. It records a digital signal at the 18.2 Mbps information rate of compressed DigiCipher. It uses two-hour metallized-tape cassettes, similar in format to 8-mm NTSC cassettes.

The proponent reports running simulations, showing that a full set of trick mode features can be supported. The DigiCipher VCR uses PCM refresh data from each field, and attempts to use DPCM data too.

Switching between compressed video images should be done at frame sync, preferably with the new scene either at black or at a scene change when the image is being processed in the PCM mode. Switching within a frame may be done at the macroblock level, with some restrictions. Otherwise, editing during frames requires decompression and recompression, with a small loss in quality do to concatenation. However, it is anticipated that most editing will be done prior to compression to the transmission rate.

No provisions have been reported for special effects in the system tested.

III. Extensibility

A. to no visible artifacts

The proponent reports simulating compression at 30 Mbps with favorable results. He believes that the algorithm can be extended to 40-45 Mbps, constituting a distribution level of quality suitable for network feeds to local affiliates. The proponent has announced that he is investigating an approach which would allow the transmission-level signal to be included in the distribution level signal as a kernel. This would permit pass-through of the transmission-level signal at the local affiliate level by stripping away the distribution-level augmentation.

B. to studio-quality data rate

The proponent speculates that studio-quality intraframe compression can be achieved at a bit rate in the 100-200 Mbps range. This format has not been developed yet.

C. to higher resolution

The proponent believes that DigiCipher technology is extensible and suggests a resolution increase by a factor of about four.

D. Provision for future compression enhancement

The proponent suggests that as decreasing digital processing costs enable increasing complexity at the encoder, improvements can be made without changing the receivers or the transmitted bit rate. These improvements are in forward analysis, perceptual analysis, motion compensation, coefficient quantization, and special effects editing.

Digital Spectrum Compatible HDTV Interoperability Assessment

System Description

The Digital Spectrum Compatible HDTV System is an all-digital system that is progressively-scanned with 787.5 lines per frame and 59.94 frames per second. The display format has square pixels with 720 lines by 1280 pixels per line in a 16:9 image aspect ratio. A feature of the system is dual-rate coding where each data segment is assigned a priority. The most important picture information on a scene-by-scene basis is transmitted as two-level vestigial sideband (2-VSB) while the remaining picture information is transmitted as 4-VSB. This allows a degree of graceful degradation on the edge of the broadcast contour and transmission robustness within the service area. It also allows a degree of bit-stream scalability. It implies a picture-dependent variable information rate, achieved with a constant symbol rate. A pilot carrier is used for easier signal acquisition because the pilot carrier is in a part of the band where co-channel NTSC receivers are not sensitive.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The proponent claims the hardware implementation of a 16-VSB transmission format to achieve a 43 Mbps data rate, which can deliver two DSC-HDTV programs over a single cable channel without any perceptible increase in the cost of terrestrial broadcast receivers.

Encryption and addressing are important service features for Cable TV. See Section II, items F and G.

B. with digital technology

Since this system is all-digital, the advantages of all-digital systems apply.

C. Headers/descriptors

The prototype tested did not have explicit headers and descriptors. However, ancillary data space was provided for a number of purposes including headers/descriptors. A data format is presently under development that includes a minimum of fixed data. Within that fixed data is a location system header that specifies the use of the remaining data. The use of data is flexible, and the header deals with error correction and priority transmission of the data. Detailed information was provided to the Review Board.

D. with NTSC

As the Digital Spectrum Compatible HDTV system is directly related to NTSC, transcoding to NTSC is straightforward. Conversion to and from NTSC has been demonstrated using real-time hardware. Edge-crop conversion to CCIR 601 involves discarding 160 HD samples from each end of the horizontal line and 4:3 interpolation of the remaining samples. It can involve 3:1 and 2:1 vertical

interpolations in alternate fields. For letter-box, 16:9 horizontal interpolation is required, as well as 3:1 vertical interpolation of the active lines. For edge-crop display, the 480 lines can be placed in a specific set of the 483 active lines. The first (odd) field can be displayed on lines 21, 23, ..., 499, and the second (even) field is displayed on lines 22, 24, ..., 500. The active lines 501 and 503 of the odd fields and 502 of the even fields can be displayed as black. For letter-box display, the vertical interpolation is more complicated. Up-conversion from NTSC requires line tripling, horizontal line-rate conversion and interpolation.

E. with film

The encoder buffer control automatically detects the presence of 24-frame/second or 30-frame/second scene material from film sources. When a film source is detected, an alternate buffer control algorithm is used which takes advantage of repeated frames in the source and minimizes variations in distortion between repeated frames. If film is detected, all video segments undergo two-level-mode transmission for maximum coverage area and minimum video data rate. The current implementation of the alternate buffer control for film mode was not completed in time for testing at the ATTC. However, it has since been completed and demonstrated.

F. with computers

Progressive scanning and square pixels, both used in this system, are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production, and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Also, scan conversion between interlaced and progressive systems can produce undesirable artifacts. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations. It also helps format conversion and extension to higher resolution.

G. with satellites

Digital signals with bit rates used by the proponents for terrestrial transmission can generally be accommodated in the available bandwidth of a satellite transponder. The maximum total data rate for the DSC-HDTV system in the mode tested by ATTC is 21.5 Mbps. As satellite data communication channels use a constant bit rate, the variable bit rate used by DSC for terrestrial transmission makes it necessary for the bit stream to be reformatted for satellite transmission. The reformatted bit stream must contain the data needed to permit reconstruction of the variable-rate bit stream for VSB-2/4 terrestrial modulation. Nevertheless, DSC-HDTV has the lowest data rate of any of the all-digital proponents. The proponent has suggested two-channel TDM and two-channel SCPC in a 36-Mhz transponder and two and one channel DBS scenarios.

H. with packet networks

The DSC symbols are organized in data segments that resemble packets. The segments make up data frames of duration 1/59.94 sec. In order to carry DSC on an ATM network, the data in data frames would be encapsulated in the ATM cell structure. However, circuit-switched networks use constant bit rate, while the number of bits in a data frame varies because of the two-level transmission. The proponent suggests repeating two-level segments for added robustness to fill out the data stream

for a constant-bit-rate channel. For a packet network, packets can be used as needed to carry the lesser actual bit rate.

When cell loss is detected, the decoder performs error concealment by replacing missing segments with default data or with pixel data from a previous frame. The error concealment algorithm was only partially implemented for ATTC testing, but has since been fully implemented.

I. with interactive systems

The picture achieves full quality within a few frames after a scene change. Full quality is obtained within one second after a channel change.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. Frame delays are required at the transmitter and at the receivers for coding and decoding. Rate buffers at the transmitter and at the receiver cause additional delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

The proponent claims that the delay through the encoder and decoder for the DSC-HDTV system is about 14 frames (224 ms). An enhancement to the current system allows the latency to be determined by the encoder, in case lower latency is required for interactive applications.

J. Format conversion

1. with 1125/60

Up-converting to the Common Image Format (1920 x 1080) is easily done by 2:3 interpolation horizontally and vertically. SMPTE 240M currently uses 1035 active lines. Colorimetry is SMPTE 240M.

2. with 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Present methods that involve frame dropping lead to jerky motion, but other techniques can produce acceptable images under most conditions. This difficult conversion is somewhat easier with a progressive system such as DSC than with an interlaced system.

3. with CCIR 601/60

The proponent claims that conversion to and from CCIR 601/60 has been demonstrated with excellent results. Commercially available line doublers were used to convert the CCIR 601 to progressive form, followed by a spatial interpolation.

4. with MPEG

Although the DSC-HDTV decoder shares many commonalties with MPEG1 decoders, the DSC-HDTV decoder would require modification to decode MPEG1.

5. with still image

The proponent suggests that conversions with JPEG, Photo CD and CDI are possible with straightforward spatial filtering after decompression, without the temporal effects that might be

introduced in an interlaced TV format. In simple cases, line and sample doubling or sub-sampling will suffice.

K. Scalability

It is possible to process the two-level data only and display the images corresponding to that portion of the video information. Film scenes and many normal video scenes will use almost entirely two-level coding for video, so the two-level data will result in virtually full-quality video. The decoding of only the two-level data may result in a reduced-quality image for scenes that are difficult to encode and that therefore require substantial four-level data. In such a case, the loss of the four-level data affects both the temporal and spatial resolution. The error-concealment algorithm replaces lost slices with information from a previous frame. Another spatial scaling method would produce a lower resolution picture by extracting a lowest frequency subset of the DCT coefficients from each compressed slice. Extraction of only the DC coefficients would produce a 160H x 90V pixel image. Scalability by extracting DC (or other low-order) DCT coefficients can be implemented in any proposed digital system. Where temporal scaling is needed, the process is simplified by the progressive scan used in DSC. The proponent has suggested using the motion vectors available at the decoder to perform motion-compensated frame interpolation.

The proponent suggests that picture-in-picture be done by windowing on slice (64H x 48V pixels) boundaries.

II. Scope of services and features

A. Initial Use of Ancillary Data

The proponent has provided charts showing a way in which flexibly allocated data may be used by means of headers. See Section I.C, Headers/Descriptors.

B. Audio

The DSC-HDTV system provides four independent 125,874 bps audio channels multiplexed with the video data. The prototype hardware used the Dolby AC-2 compression equipment. The main stereo pair is transmitted as two-level VSB, whereas the secondary pair is transmitted as four-level VSB. The two-level data has a 6 dB lower threshold at the fringe area.

The newly proposed use of headers provides a flexible composite surround-sound method, described in ATSC document T3/186, in place of the original fixed allocation. It also provides a separate stereo pair suitable for a second language.

C. Data

Two separate channels were provided for ancillary data in the system tested. The total capacity of 412,867 bps was divided into one channel of 30,210 bps sent as 2-level data and another of 382,657 bps sent as 4-level data. However, the newly proposed header usage allows a much more flexible assignment of ancillary data capability. Assuming the presence of video, five-channel audio, and captioning, and that all the ancillary data is carried as four-level data, the ancillary data rate is 493,426.57 bps.

D. Text

The technical description dated 2/22/91 suggested that 35 bytes per data field or 16,783 bps be used for teletext. The new proposal would flexibly allocate text data by use of headers and descriptors.

E. Captioning

The technical description dated 2/22/91 suggested that 1 byte per data field or 480 bps be used for captioning. The new proposal would allocate 9.6 kbps per ATSC Document T3/186.

F. Encryption

The prototype tested did not have encryption implemented. The proponent expects to develop encryption with industry participation.

G. Addressing

The use of headers under development allows program quality to be traded temporarily for addressing performance. An extreme mode of the system in which most of the data rate is dedicated to addressing would allow addressing of more than 5.7 million subscribers per minute. In this mode, other auxiliary data would be suppressed, and a minimum quality of audio/video service would be maintained.

H. Low-cost receiver

The proponent described an algorithm for producing a 640H x 360V pixel image by extracting one out of four 8x8 blocks of pixels and tiling them. The resulting picture is said to exceed the quality of NTSC on a 2" LCD display. It is also reported possible to make a low-cost receiver with just the two-level data.

I. VCR capability

The proponent claims that the current S-VHS mechanism is sufficient for the 21.5 Mbps compressed DSC-HDTV data, and that such 1/2 inch cassette equipment exists in prototype form.

According to the proponent, the system knows what fraction of the original image is contained in the displaced frame difference (DFD). A usable picture is obtained without motion compensation by amplifying the DFD by a factor proportional to the inverse of the DF factor. This can be used for VCR forward or reverse scan modes when only a small portion of each compressed frame is acquired. In addition, the segment headers are needed to identify the slice numbers from the acquired data. The picture would appear "blocky" with some slices lost, but suitable for rapid searching.

Still frame is simple if the VCR had been playing. If random access to a particular frame on the tape is required, the decoding of several frames leading up to it is needed to achieve full quality.

Splicing is optimal if each splice starts with a scene change. Otherwise, the decoder can be signaled to initiate a leak factor inversion for fast startup at the beginning of each splice or insert.

Cropping is possible by manipulation or replacement of compressed slices.

Image processing for special effects is best performed in the pixel domain after decoding. Square pixels and progressive scanning simplify the implementation of special effects.

III. Extensibility

A. to no visible artifacts

It was the reported intention of this proponent to have no visible artifacts at the terrestrial broadcast data rate, when viewed from a distance of 3 times the picture height. This goal is, however, challenged by complexity and motion. The proponent suggests a rate of 41 Mbps for no visible artifacts

regardless of detail and motion, and states that this can be accomplished with a small change to the compressed video interface.

B. to studio-quality data rate

Claims are made that the compression techniques used for the broadcast of DSC-HDTV are easily simplified to produce a 200 Mbps signal for use in the studio. This signal uses only intraframe processing, and thus is suitable for all editing and special effects processing. The claim is made that the quality may be suitable for multiple decoding/encoding as required. The bit rate is suitable for serial data interfaces and also for video tape recording on D-1 VTRs.

C. to higher resolution

If it is desirable in the future to maintain higher pixel numbers in the production studio, this can also be accommodated in an extension of the studio compression format, by compressing the higher-resolution signal rate into the 200 Mbps studio compression-standard plus a high-frequency residual signal. The standard DSC-HDTV system codes the HD frames, and a simple augmentation encoder codes the residual signal. The final output of editing or special effects can still be recorded using the 200 Mbps portion of the compressed signal.

D. Provision for future compression enhancement

The compression algorithm permits improvements in the selection of vector quantization patterns, motion estimation, perceptual error threshold computation, buffer control, leak adaptation, and transmission prioritization. These improvements can be made without change to the receivers or the transmitted data rate.

Narrow-MUSE Interoperability Assessment

System Description

The Narrow-MUSE system is an analog transmission simulcast system. The input signal format and the display signal format is 1125 scanning lines per frame, 60.00 fields per second, and 2:1 interlace. The transmission format, however, is 750 scan lines per frame, 60.00 fields per second, and 2:1 interlace. The video encoding method is Multiple Sub-Nyquist Sampling Encoding, and the audio encoding method is near-instantaneous DPCM. Transmission is by analog amplitude modulation.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 Mhz for improved service. MUSE-E of 8.1 Mhz may satisfy this criterion. An even more desirable feature for a system to offer would be a higher-quality mode within the 6-Mhz channel. Digital systems can do this by sending more bits per symbol, for use over quiet channels such as cable. Since the Narrow-MUSE system is an analog system, it cannot be enhanced in this manner.

Encryption and addressing are other important service features for cable TV. See Section II, items F and G.

B. with digital technology

Since the transmitted signal is analog, it must be digitized before interfacing with digital technology. However, all of the signal processing in the encoder, modulator, demodulator and decoder is done in the digital domain, and custom LSI's are available. A digital interface port is provided in the receiver.

C. Headers/descriptors

This system provides 128 kbps of ancillary data. Although the partition of the ancillary data has not yet been specified, the proponent states that the headers/descriptors as a part of SMPTE 240M and 260M could be assigned into the ancillary data channel of Narrow-MUSE.

D. with NTSC

There are two conversion methods to NTSC. One is from the 750-line Narrow-MUSE transmission format to NTSC, and the other is from 1125/60 to NTSC. Since the aspect ratio of this system, 16:9, is different from the aspect ratio of NTSC, 4:3, a choice must be made for the conversion. The limiting choices for downconversion are 1) "Edge Crop", in which the HDTV picture fills 483 lines on NTSC with loss of the sides of the HDTV picture, and 2) "Letterbox", in which the full width of the HDTV picture is displayed in the full width of the NTSC picture, but leaving unused areas at the top and the bottom of the NTSC screen.

The conversion from Narrow-MUSE to NTSC requires only vertical interpolation because Narrow-MUSE employs an analog transmission technique. Since the number of active lines in Narrow-MUSE is 651, the vertical interpolation ratio is 31:23 for "Edge-Crop" and is 9:5 for the letterbox format.

The interpolation ratio for 1125 to NTSC is 15:7 vertically and 2:1 horizontally. In both conversions, field-rate conversion from 60.00 Hz to 59.94 Hz is required. The proponent claims that a motion-adaptive field-rate converter is available, and has been publicly demonstrated. The proponent also claims that the converter for 1125/60 to NTSC is used for the daily simulcast operation in Japan, that the converter from full-band MUSE to home display is sold on the market, and that the same technique can be applied to Narrow-MUSE.

E. with film

This system does not have a film mode within its encoding algorithm. Since the field rate of this system is 60 Hz, the temporal conversion from film to HDTV is accurate. A motion-compensated continuous-film-transfer telecine is already available for this system.

F. with computers

For computers, a progressive display is overwhelmingly preferred. Interlaced systems such as this have an inherent compatibility problem with computers, even for just displaying the computer's video output. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced displays of computer-generated imagery. Also, scan conversion between interlaced and progressive systems can produce undesirable artifacts. Square pixels are also overwhelmingly preferred for graphics display hardware for all computers employing bit-mapped displays, as it facilitates processing of 2D transformations, especially rotations. Non-square pixels do not present a problem for the display of RGB signals from a computer's video card on HDTV receivers, but would complicate more sophisticated attempts at interoperability, such as the display by an HDTV receiver of a picture from a digital data stream generated by a computer. According to the proponent, the pixel shape of 240M is 1 : 1.043 (H : V), and 1125/60 signals have already been manipulated for graphic purposes. The shape of a decoded Narrow-MUSE pixel is 1 : 0.56. The proponent claims that the field rate of 60.00 Hz is a better selection than 59.94 Hz for interoperability with computers that have integer field rates such as 72 Hz.

G. with satellites

This system can be transmitted through a satellite using FM with an RF channel bandwidth of approximately 15 MHz. FM transmission of MUSE through a satellite is a proven technology. The proponent claims that Narrow-MUSE also can be transmitted through a satellite using digital transmission. The Narrow-MUSE signal can be encoded by DPCM to a data rate of approximately 40 Mbps which includes error correction of 8% (3.2 Mbps). Satellite links typically use more error correction than this, e.g. 14% to 50%. The RF channel bandwidth with QPSK is approximately 24 MHz. Digital transmission of MUSE in conjunction with DPCM is also a proven technology.

H. with packet networks

This item is not applicable because this system employs analog transmission.

I. with interactive systems

The acquisition time is 0.5 seconds.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. Frame delays are required at the transmitter and at the receivers for coding and decoding. Because the digital systems have other delays, comparison of latency between the proponent systems is difficult at this time.

According to the proponent, the total delay for Narrow-MUSE through an encoder and a decoder is 6 fields (approximately 100 msec), 3 fields for each.

J. Format conversion

1. with 1125/60

No format conversion is required because this system uses 1125/60 format (SMPTE 240M). The decoded Narrow-MUSE signal can be converted to the Common Image Format through a line-number conversion and a sampling-frequency conversion. These are 24:23 vertically and 50:27 horizontally.

2. with 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. However, this system makes the frame-rate conversion slightly easier than other systems, because its field rate is exactly 60.00 Hz. The proponent claims that the conversion from 60 Hz to 50 Hz has been successfully demonstrated with a standard converter from 1125/60 to PAL. This standard converter employed interpolation based on motion vectors. The vertical interpolation ratio from 1125 to 1250 is 9:10.

3. with CCIR 601/60

The vertical interpolation ratio between this system and CCIR 601/60 is 15:7 for both total scanning lines (1125:525) and active scanning lines (1035:483). The horizontal conversion ratio is 2:1. The clock frequency for this system, 74.25 MHz, and for CCIR 601, 13.5 MHz, have a relationship of 11:2.

Because the field frequency for this system is 60.00 Hz, field-rate conversion from 60.00 Hz to 59.94 Hz is required. A motion-adaptive field-rate converter is available and is used daily for simulcast in Japan. A frame skip must take place every 33 seconds, because of the 1001/1000 frame conversion. The hardware attempts to do this cut on a motionless picture or on a scene change.

4. with CCIR 601/50

The vertical conversion ratio between Narrow-MUSE and CCIR 601/50 is 9:5 for both total scanning lines (1125:625) and active scanning lines (1035:575). The horizontal ratio is 2:1. The clock frequency for Narrow-MUSE, 74.25 MHz and for CCIR 601, 13.5 MHz, have a relationship of 11:2. Also, see item 2 above.

5. with MPEG

Narrow-MUSE does not have interoperability with MPEG.

6. with still image

An 1125/60 still image disk system based on the JPEG algorithm has been developed and demonstrated. The system uses multiple disks, with video, audio, and control data recorded on separate disks. Narrow-MUSE does not have compatibility with JPEG, Photo CD, or CDI.

K. Scalability

This system uses a multiple sub-sampling technique with a four-field sequence. Therefore, the spatial resolution of the reconstructed picture can be controlled by selecting fields to be used for the interpolation. When all four fields are used, a full-quality picture is obtained. When one of every four fields is used, a picture with reduced resolution can be obtained by interpolation. Also, a picture with reduced size can be obtained by using only a selected field.

The proponent claims that Narrow-MUSE is a member of the MUSE family, which is based on the concept of scalability. The MUSE family consists of MUSE-T, MUSE-E (full-band MUSE), Narrow-MUSE, and NTSC MUSE-4, all based on the same coding algorithm. All these systems have been demonstrated.

For display on a computer, pictures reduced by $1/2^n$ can be made with only intra-field information. Other ratios require more processing.

This system uses a multiple sub-sampling technique with a four-field sequence. Therefore, the temporal resolution of the reconstructed picture can be controlled by selecting fields to be used for the interpolation. When all four fields are used, full temporal resolution, 1/60 sec, is obtained. To reduce the amount of data, the field repetition rate can be reduced for pictures with less temporal resolution.

The multiple sub-sampling technique makes possible two types of receivers differing in complexity. A simple receiver can be built that handles only intra-field interpolation, while the full-capability receiver handles both intra-field and inter-field interpolation.

The low-frequency component below 2 MHz of the Narrow-MUSE signal does not contain the aliasing component caused by frame offset sub-sampling. Therefore, a picture whose quality is equivalent to NTSC can be reproduced by using only this low-frequency component.

Picture-in-picture, picture-out-of-picture, and multiple programs can be accommodated using only the intra-field information from the Narrow-MUSE signal. A frame store in the receiver can be used for this purpose.

II. Scope of services and features

A. Initial use for ancillary data

The partition of the ancillary data has not been specified yet.

B. Audio

This system provides two audio modes. Both modes employs Near-instantaneous Companding DPCM for coding. Mode A can transmit four audio channels with 15-kHz bandwidth, the data rate of which is 1056 kbps/channel without error protection. Mode B can transmit two audio channels with 20-kHz bandwidth, the data rate of which is 1072 kbps/channel without error protection. Since the data capacity of the audio channel is relatively high, modifications can be made to increase the number of audio channels.

C. Data

This system provides 128 kbps of ancillary data. This can be increased by modifying the audio encoding method. The interface for the data channel is RS-422.

D. Text

Teletext data is transmitted using the ancillary data channel. The proponent suggests a data rate of 16 kbps. He also suggests that the number of characters per picture is between 1000 and 4000, to be transmitted in about 3 seconds. These numbers can vary, based on the service required.

E. Captioning

Captioning data is transmitted using the ancillary data channel. The proponent suggests a data rate of 0.6 kbps. This number might vary, based on the service required.

F. Encryption

The system submitted for testing did not include encryption. The proponent suggests a combination of line rotation and line permutation for signal security. Decoder chips for the encryption system are already developed.

G. Addressing

The proponent claims that the data rate needed for addressing and sending decryption keys is approximately 16 kbps, and that this rate will permit the addressing of 10,000,000 subscribers in approximately two days. However, this addressing rate is not sufficient for a pay-per-view environment. The addressing information is transmitted through the ancillary data channel.

H. Low-cost receiver

This system has two modes in the receiver, i.e. stationary mode and motion mode. The stationary mode requires more memory and a more complicated interpolation than the motion mode. By using only the motion mode, low-cost receivers can be built with reduced resolution. The quality is higher than with NTSC on a 2" LCD receiver because the resolution is almost the same, and neither cross-luminance nor cross-color is observed. The proponent claims that a reduced-cost receiver is already available on the market for the full-band MUSE, and that the same configuration can be used for Narrow-MUSE.

I. VCR capability

The proponent claims that a digitized Narrow-MUSE signal with an 80-Mbps data rate or a DPCM-encoded Narrow-MUSE signal with a 40-Mbps data rate can be digitally recorded on a 1/2 inch cassette VCR. The record/playback time is said to be 4 to 8 hours. Since the bandwidth of Narrow-MUSE is the same as for NTSC, the requirements of a VCR are similar.

Only sync blocks whose ID signals are detected correctly are used for fast forward and reverse. Sync blocks whose ID signals are not detected correctly are discarded, and replaced with interpolated

information. The proponent claims that the quality of a rapid search picture will be comparable to that of a 4-head VHS machine.

These functions can be achieved, based on the four-field sequence of the Narrow-MUSE algorithm. Editing functions can be implemented by adjusting the subsampling phases between the materials to be edited, using the subsampling phase information which is transmitted as a part of the control signal. The four-field sequence is similar to that of NTSC.

Special effects are not done with the Narrow-MUSE signal. Rather, they are done on the 1125/60 signal.

III. Extensibility

A. to data rate with no visible artifacts

MUSE-T, a higher member of the MUSE family, has a bandwidth of 16.2 MHz and can provide a picture with no visible artifacts because it employs only intra-field subsampling. A digitized MUSE-T can be further compressed using DPCM without introducing additional artifacts. The main part of a Narrow-MUSE receiver can be shared for MUSE-T decoding when MUSE-T is transmitted through alternate media such as DBS. It is also possible to extend Narrow-MUSE to MUSE-T by transmitting the difference between locally decoded Narrow-MUSE and MUSE-T through an additional channel.

B. to studio quality data

It is possible to extend Narrow-MUSE to 240M by transmitting the difference between the locally decoded Narrow-MUSE and 240M signals through an additional channel as augmentation information. The bandwidth of the studio-quality signal is 60 MHz (30 MHz for luminance signal and 15 MHz for each color difference signal).

C. higher resolution

MUSE-E can be used for higher resolution. The main part of a Narrow-MUSE receiver can be shared for MUSE-E when MUSE-E is transmitted through alternate media such as DBS. It is also possible to extend Narrow-MUSE to MUSE-E by transmitting the difference between locally decoded Narrow-MUSE and MUSE-E through an additional channel.

The proponent suggests that it is possible to extend Narrow-MUSE to VHDTV and UHDTV by transmitting the difference between the locally decoded Narrow-MUSE and VHDTV/UHDTV through an additional channel as an augmentation signal.

D. Provision for future compression enhancement

The proponent claims that the dynamic resolution can be improved by increasing the number of motion vectors. The additional motion vectors can be transmitted through the data channel at the expense of data for other purposes.

APPENDIX B

ANALYSIS OF INTEROPERABILITY EVALUATIONS BY PROPONENT AND REVIEW BOARD

Reference 2

ANALYSIS OF INTEROPERABILITY EVALUATIONS BY PROPONENTS AND REVIEW BOARD

DRAFT - November 16, 1992

StellaCom, Inc.
Advanced Television Systems
1901 N. Moore Street - Suite 601
Rosslyn, VA 22209-1706

(703) 243-1695 (p)
(703) 243-1698 (f)

This work was done for ACATS PS/WP-4. Financial support was provided by Apple Computer, Inc., Bellcore, Cable Television Laboratories, Inc., Eastman Kodak Company, Digital Equipment Corporation, and Viacom International, Inc.

INTRODUCTION

Following interoperability presentations made on September 23-25, 1992 by the advanced television system proponents, the Review Board of multi-industry experts and the system proponents were asked to evaluate the ATV systems with regard to interoperability, extensibility, and scope of services and features using the ATV Interoperability Evaluation Matrix previously defined. An analysis of the evaluation matrices submitted in response to this request follows in this report.

The interoperability assessment was a qualitative assessment process. The definitions of the "ATV Features" in the evaluation matrix are not precise and while "generally understood" by the Review Board, differences in understanding of the features exists.

It was difficult to delineate (1) proponent system configuration tested (2) proponent system configuration that would be field tested (3) proponent system configuration that would be first delivered (4) possible/realistic future configurations for a proponent system. As much as possible, the evaluations were based on "what the evaluator expects the proponent will deliver commercially".

The weighting of the value (e.g. , market/social benefit) of interoperability or extensibility other than in established markets (e.g. terrestrial broadcast or cable television, entertainment, advertising and news) is difficult and arguable. Formal positions from the affected industries are not available.

In response to the interoperability investigation, some proponents suggested alternative configurations from that tested by the ATTC for improving interoperability. Proponents were invited to submit additional information on interoperability by October 1, 1992. Two proponents submitted such information on that date and this was incorporated into the draft assessment dated October 19, 1992. Since the evaluation matrices were due on October 5, the additional proponent information may not have been considered in some of the evaluations. For example, a Review Board member indicated that his rating for DigiCipher packet network interoperability should change from "4" to "2" with the implementation of the video packet and transwitch format layers described in the October 1, 1992 DigiCipher supplemental information. Review Board members were invited to submit any changes they wished to make in their evaluations by November 2, 1992. The revised evaluations are included in Appendix III.

Proponent Analysis of Other Systems

In general, each proponent evaluated his own system better than his opponents (colleagues) evaluated his system. In fact, one proponent did not include the evaluation of his own system because of "possible conflict of interest". Matrix I, on the next page, was done without the proponent's evaluation of his own system to remove this bias. Only two proponents evaluated interoperability with cable TV and hence this was not included. The ratings of all proponents are found in Appendix I.

The mean value found in the tables is the estimate of the mean based on N independent observed values, x_i, of the random variable x.

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

The standard deviation in the tables is the unbiased estimate given below.

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Interoperability

with Digital Technology

The all-digital systems were rated approximately equal. The Narrow-MUSE system was evaluated as difficult to interoperate with digital technology. The proponent did not agree with this assessment.

with Headers/Descriptors

AD-HDTV was rated as easiest to implement with headers/descriptors and Narrow-MUSE the hardest with no disagreement from the Narrow-MUSE proponent.

MATRIX I - PROPONENT ANALYSIS OF OTHER SYSTEMS*

ATV Interoperability Evaluation Matrix

Mean (Standard Deviation)

ATV FEATURE	ATV Proponent System				
	Advanced Digital	CCDC	DigiCipher	DSC	Narrow-Muse
Interoperability					
with digital technology	1.25 (0.50)	1.25 (0.50)	1.00 (0.00)	1.25 (0.50)	4.75 (0.50)
with headers/descriptors	1.25 (0.50)	2.50 (1.00)	2.25 (0.96)	2.25 (0.96)	5.00 (0.00)
with NTSC	1.75 (0.50)	2.00 (1.15)	1.50 (0.58)	2.25 (0.96)	2.63 (1.11)
with film	1.75 (0.50)	1.25 (0.50)	1.25 (0.50)	2.00 (0.82)	3.75 (0.96)
with computers	2.75 (0.50)	2.25 (0.96)	3.25 (0.96)	2.25 (0.96)	5.00 (0.00)
with satellites	1.75 (0.96)	1.50 (0.58)	1.50 (0.58)	2.25 (0.96)	2.50 (0.58)
with packet networks	1.25 (0.50)	2.75 (0.50)	2.75 (0.50)	2.75 (0.50)	5.00 (0.00)
with interactive systems	2.75 (0.50)	2.25 (0.96)	2.75 (0.50)	2.75 (0.50)	4.50 (1.00)
Format conversion	2.75 (0.50)	2.00 (0.82)	2.75 (0.50)	2.25 (0.50)	4.00 (0.82)
Scalability	2.50 (1.29)	3.00 (0.82)	3.50 (1.00)	2.50 (1.00)	4.50 (1.00)
Scope of Services and Features					
Initial use for ancillary data	2.00 (1.15)	2.00 (1.15)	2.00 (1.15)	2.00 (1.15)	2.75 (1.71)
Audio	1.50 (1.00)	1.38 (0.75)	1.00 (0.00)	1.25 (0.50)	1.75 (0.96)
Data	1.50 (1.00)	2.50 (1.00)	2.00 (1.15)	2.00 (1.15)	3.38 (1.25)
Text	1.50 (1.00)	2.25 (0.96)	1.75 (0.96)	2.00 (1.15)	3.13 (1.44)
Captioning	1.50 (1.00)	2.25 (0.96)	1.75 (0.96)	2.00 (1.15)	3.00 (1.41)
Encryption	1.75 (0.96)	1.50 (1.00)	1.25 (0.50)	1.75 (0.96)	3.75 (1.26)
Addressing	1.75 (0.96)	1.75 (0.96)	2.00 (1.15)	2.50 (1.00)	3.50 (0.58)
Low-cost receiver	2.63 (0.95)	2.75 (0.50)	2.50 (0.58)	2.13 (0.25)	2.25 (0.96)
VCR capability	2.25 (0.50)	2.88 (0.85)	2.75 (0.96)	2.88 (1.03)	2.75 (0.96)
Extensibility					
to no visible artifacts	2.25 (0.96)	2.38 (0.48)	2.25 (0.96)	2.25 (0.96)	4.00 (0.82)
to studio-quality data rate	2.25 (0.96)	2.50 (0.58)	2.25 (0.96)	2.25 (0.96)	3.63 (1.11)
to higher resolution	2.38 (0.48)	2.75 (0.50)	2.75 (0.96)	2.50 (0.58)	4.25 (1.50)
to VHDTV	3.00 (1.42)	3.50 (0.58)	3.25 (0.96)	3.25 (0.96)	4.63 (0.75)
to UHDTV	3.00 (1.42)	3.50 (0.58)	3.25 (0.96)	3.25 (0.96)	4.63 (0.75)
Provision for future compression enhancement	2.50 (0.58)	2.75 (0.50)	2.75 (0.50)	2.75 (0.50)	4.25 (0.96)

*Proponent rating of own system not included.

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

with NTSC

All systems were evaluated as relatively easy to interoperate with NTSC. DigiCipher received a slight edge here.

with Film

DigiCipher and CC-DigiCipher received a slight edge in interoperability with film.

with Computers

DSC-HDTV and CC-DigiCipher (progressive scan/square pixels) received a slight edge in interoperability with computers.

with Satellites

All systems were judged as relatively easy to interoperate with satellites with DigiCipher and CC-DigiCipher having a slight edge.

with Packet Networks

AD-HDTV was judged significantly easier ($>2\sigma$) to interoperate with packet networks.

with Interactive Systems

The all-digital systems were judged in the same range of difficulty in interoperability with interactive systems, however, CCDC had a slight edge.

Format Conversion

The progressive scan systems achieved a slight edge in ease of format conversion with CCDC having a slight edge.

Scalability

AD-HDTV and DSC-HDTV received a slight edge in scalability.

Scope of Service and Features

Initial use for ancillary data

No significant difference.

Audio

No significant difference.

Data
Text
Captioning

AD-HDTV had a slight edge in data services and features but the other systems were not far behind.

Encryption

DigiCipher had a slight edge in encryption.

Addressing

No significant difference in addressing capability was shown among all-digital systems although AD-HDTV and CC-DigiCipher received a slight edge.

Low-cost receiver

No significant difference but DSC-HDTV had a very slight edge.

VCR capability

AD-HDTV was given an edge in VCR capability.

Extensibility

to no visible artifacts
to studio-quality

No significant difference among all-digital systems.

to higher resolution
to VHDTV, UHDTV

No significant difference among all-digital systems with AD-HDTV having only the slightest edge.

Provision for Future Compression Enhancement

No significant difference among all-digital systems with AD-HDTV having only the slightest edge.

Summary Evaluation

Interoperability

No one system is best for interoperability with all media. In order to arrive at some conclusion, assume that all features under Interoperability are weighted uniformly. Then the following Interoperability table is computed with mean (standard deviation).

AD-HDTV	CCDC	DigiCipher	DSC-HDTV	Narrow-MUSE
1.98(0.86)	2.08(0.92)	2.25(1.03)	2.25(0.81)	4.16(1.11)

Only very slight differences are noted for the all-digital systems. AD-HDTV edged out CCDC, which edged out the other two all-digital systems.

Scope of Services and Features

It is not advisable to weight all features equally. Audio, VCR capability, addressing and low-cost receiver are arguably more important. Different systems lead the rankings in each category. Therefore, no conclusion can be reached for this category.

Extensibility

Assuming that all features under Extensibility are weighted uniformly, the following Extensibility table is computed with mean (standard deviation).

AD-HDTV	CCDC	DigiCipher	DSC-HDTV	Narrow-MUSE
2.52(1.02)	2.90(0.66)	2.75(0.90)	2.71(0.86)	4.23(0.97)

Only slight differences are noted for the all-digital systems. AD-HDTV edged out DSC-HDTV, which edged out DigiCipher. AD-HDTV ranked best or tied for best for all extensibility features and would, therefore, be ranked best or tied for best regardless of the weighting of the features.

II. Evaluation by the Review Board

A tabulation of the means and standard deviations from the evaluations by the Review Board is shown as Matrix II. The evaluations by all Review Board members are given in Appendix II. In many cases the information was incomplete (n) or verification (v) was needed. The number of times that "n" and "v" appeared in the review board ratings is noted by the number inside the parentheses in the matrix that follows.

Interoperability

All five systems were judged to be easily deliverable by satellite or Cable TV. AD-HDTV was rated most interoperable with headers/descriptors and packet networks. The four all-digital systems were judged to interoperate well with film with CCDC ranking better than the other three systems. Rankings of other characteristics of interoperability were close. DSC-HDTV was most interoperable with computers with CCDC close, and best for format conversion with AD-HDTV close. DSC-HDTV was judged most interoperable with digital technology and interactive systems, with other all-digital systems very close. CCDC was ranked most interoperable with NTSC, with DSC-HDTV very close. AD-HDTV was ranked most scalable with DSC-HDTV very close. DigiCipher was ranked slightly more interoperable with cable TV, with all other systems very close. If all features are weighted uniformly, the overall mean for AD-HDTV is 1.88, DSC-HDTV is 1.90, CCDC is 2.13, DigiCipher is 2.43 and Narrow-MUSE is 3.61.

Scope of Features and Services

The spread between evaluations of the various systems was small. However, AD-HDTV ranked best in initial use for ancillary data, data, encryption, addressing, low-cost receiver and VCR capability. DSC-HDTV ranked best in text and captioning. CCDC ranked best in audio. Unless substantially higher weighting is placed on the last three features, AD-HDTV would be ranked best for scope of services and features.

Extensibility

No significant differences are noticed among the mean values of the all-digital systems. If it is assumed that all features under Extensibility are weighted uniformly, DSC-HDTV has the lowest mean (best score) with 2.43, AD-HDTV and CCDC are tied with a mean of 2.53, DigiCipher has a mean of 3.08 and Narrow-MUSE of 4.50.

Matrix II - REVIEW BOARD Evaluation

ATV Interoperability Evaluation Matrix - Revision A
mean (standard deviation)

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	DigiCipher	DSC	Narrow-Muse	Footnotes
Interoperability						
with cable TV	1.22 (0.44)	1.25 (0.46)	1.13 (0.35)	1.25 (0.46)	1.89 (1.45)	
with digital technology	1.67 (0.84)	1.40 (0.57)	1.61 (1.09)	1.32 (0.52)	3.97 (1.00)	
with headers/descriptors	1.36 (0.50)	2.73 (0.79)	2.90 (0.99)	2.00 (0.94)	4.50 (0.74)	
with NTSC	1.91 (0.94)	1.45 (0.69)	1.80 (0.92)	1.60 (0.84)	3.05 (1.15)	
with film	1.55 (0.69)	1.18 (0.40)	1.60 (0.84)	1.60 (0.84)	3.59 (1.16)	
with computers	2.49 (0.92)	2.04 (0.78)	3.51 (0.84)	1.83 (0.65)	4.38 (0.93)	
with satellites	1.18 (0.40)	1.27 (0.47)	1.20 (0.42)	1.20 (0.42)	1.86 (1.00)	
with packet networks	1.27 (0.47)	3.20 (0.42)	3.20 (0.79)	2.30 (0.67)	5.00 (0.00)	
with interactive systems	2.34 (1.05)	2.38 (0.86)	2.57(1.34)	2.26 (0.83)	3.57 (1.17)	
Format conversion	2.45 (1.13)	2.64 (1.21)	3.00 (0.94)	2.20 (0.79)	4.09 (1.04)	
Scalability	3.18 (1.25)	3.84 (0.98)	4.20 (0.63)	3.30 (1.06)	3.80 (1.23)	
Scope of Services and Features						
Initial use for ancillary data	1.67 (1.00)	1.78 (0.97)	1.75 (1.04)	1.75 (1.04)	2.33 (2.00)	
Audio	1.20 (0.42)	1.09 (0.30)	1.20 (0.42)	1.20 (0.42)	1.82 (1.47)	
Data	1.36 (0.67)	1.64 (0.81)	1.70 (1.06)	1.50 (0.85)	2.27 (1.56)	
Text	1.50 (1.27)	1.40 (0.70)	1.56 (1.33)	1.22 (0.44)	2.00 (1.41)	
Captioning	1.55 (1.21)	1.36 (0.50)	1.60 (1.26)	1.30 (0.48)	2.09 (1.30)	
Encryption	1.36 (0.67)	1.45 (0.82)	1.40 (0.84)	1.40 (0.70)	2.91 (1.22)	
Addressing	1.18 (0.40)	1.68 (1.19)	1.30 (0.95)	1.40 (0.70)	2.55 (1.21)	
Low-cost receiver	2.91 (1.04)	3.30 (0.95)	3.40 (0.70)	3.00 (0.94)	3.09 (1.22)	
VCR capability	2.18 (0.60)	2.73 (0.47)	2.70 (0.48)	2.60 (0.52)	3.50 (1.18)	
Extensibility						
to no visible artifacts	2.00 (0.94)	1.70 (0.67)	2.33 (1.00)	1.89 (0.60)	4.20 (1.03)	
to studio-quality data rate	2.00 (0.94)	2.00 (0.67)	2.67 (1.12)	1.89 (0.60)	4.05 (0.90)	
to higher resolution	2.80 (1.23)	3.00 (0.94)	3.56 (0.88)	2.78 (1.09)	4.50 (0.85)	
to VHDTV	3.00 (1.05)	3.00 (0.94)	3.56 (0.88)	2.78 (1.09)	4.80 (0.63)	
to UHDTV	3.00 (1.05)	3.00 (0.94)	3.56 (0.88)	2.78 (1.09)	4.80 (0.63)	
Provision for future compression enhancement	2.40 (0.97)	2.50 (1.08)	2.78 (1.30)	2.44 (1.01)	4.65 (0.67)	

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
2	
3	- Moderately difficult
4	
5	- Very difficult to implement
n	No answer or incomplete
v	Verification needed

MATRIX II - REVIEW BOARD Evaluation (con't)

ATV Interoperability Evaluation Matrix - Revision A

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	DigiCipher	DSC	Narrow-Muse	Footnotes
Interoperability						
with cable TV						
with digital technology	v(1)	v(1) n(1)	v(1) n(1)	v(1) n(1)	v(1) n(1)	
with headers/descriptors	v(1)	v(1)	v(1)	v(1)	n(1)	
with NTSC	v(1)	v(1) n(1)	v(1) n(1)	v(1) n(1)	v(1) n(1)	
with film	v(2)	v(1)	v(1)	v(1)		
with computers	v(2)	v(2) n(1)	v(1) n(1)	v(2) n(1)	v(1)	
with satellites						
with packet networks	v(2) n(1)	V(1) n(3)	n(1)	v(1) n(1)	n(3)	
with interactive systems	v(2)	v(1) n(2)	n(1)	v(2)		
Format conversion	v(1)	v(2)	v(1)	v(2)	v(1)	
Scalability	v(1)	v(2)	v(1)	v(2)	v(1)	
Scope of Services and Features						
Initial use for ancillary data	v(1) n(1)	v(1) n(1)	v(1) n(1)	v(1) n(1)	n(1)	
Audio						
Data	v(1)	v(1)	v(1)	v(1)		
Text		v(1)		v(1)		
Captioning		v(1)		v(1)		
Encryption	v(2)	v(2)	v(2)	v(2)		
Addressing	v(2)	v(2)	v(2)	v(2)		
Low-cost receiver	v(2)	v(2) n(1)	v(2)	v(2)	v(1)	
VCR capability	v(2)	v(2)	v(1)	v(2)	v(1) n(1)	
Extensibility						
to no visible artifacts	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(1) n(2)	
to studio-quality data rate	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(1) n(2)	
to higher resolution	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(1) n(2)	
to VHDTV	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(1) n(2)	
to UHDTV	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(2) n(2)	v(1) n(2)	
Provision for future compression enhancement	v(2) n(2)	v(2) n(2)	v(1) n(2)	v(2) n(2)	v(1) n(2)	

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
2	
3	- Moderately difficult
4	
5	- Very difficult to implement
n	No answer or incomplete
v	Verification needed

III. Experts Interoperability Evaluation

The previous analyses involved averaging a number of evaluations. In the meeting on Friday, September 25, several experts from the Review Board led the discussion of interoperability issues for evaluating the advanced television systems. Each discussed features from the evaluation matrix important to the area of technology in which he is an expert. The following Experts' Evaluation Matrix consists of the evaluations by each expert of the features that he discussed.

Interoperability

All systems were judged interoperable with cable TV. The progressive-scan square-pixel systems were judged most interoperable with digital technology. AD-HDTV was judged most interoperable with headers/descriptors. The progressive scan systems were judged most interoperable with NTSC. AD-HDTV and DSC-HDTV were judged most interoperable with film. DSC-HDTV was judged most interoperable with computers. All systems were judged interoperable with satellites. DSC-HDTV was judged most interoperable with interactive systems. The progressive systems were judged best for format conversion. AD-HDTV, DSC-HDTV and Narrow-MUSE were ranked best for scalability. DSC-HDTV was ranked best or tied for best for all features except headers/descriptors. If all features were weighted equally, DSC-HDTV would be rated the most interoperable.

Scope of Features and Services

DSC-HDTV and AD-HDTV were ranked best for data and for VCR capability. DSC-HDTV was tied with Narrow-MUSE for low-cost receiver and was tied with CCDC for text interoperability. DSC-HDTV would, therefore, be given the best overall evaluation for scope of services and features.

Extensibility

AD-HDTV was rated best or tied for best in all categories and would, therefore, be most extensible.

MATRIX III- EXPERTS' EVALUATION

ATV Interoperability Evaluation Matrix - Revision A

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	DigiCipher	DSC	Narrow-Muse	Footnotes
Interoperability						
with cable TV	1	1	1	1	1	C. Tanner
with digital technology	3	2	4	2	5	G. Demos
with headers/descriptors	2	3n	3n	3n	5	M. Liebhold
with NTSC	2	1	2	1	3	A. Uyttendaele
with film	1	2	2	1	4	A. Toth
with computers	2.42v	2.42nv	3.1nv	2.28nv	3.7v	M. Liebhold
with satellites	2	2	2	2	2	A. Uyttendaele
with packet networks	2	3	4/ 2	2	n	J. Bellisio
with interactive systems	2.42v	2.42nv	3.1nv	2.28nv	3.7v	M. Liebhold
Format conversion	5	2v	4	2v	5	G. Demos
Scalability	3	4	4	3	3	J. Bellisio
Scope of Services and Features						
Initial use for ancillary data	1	1	1	1	1	J. Cohen
Audio	1	1	1	1	1	J. Cohen
Data	1	2	2	1	2	J. Cohen
Text	5	2v	5	2v	5	G. Demos
Captioning	1	1	1	1	1	G. Hanover
Encryption	1	1	1	1	3	C. Tanner
Addressing	1	1	1	1	3	C. Tanner
Low-cost receiver	3	3	3	2	2	G. Hanover
VCR capability	2/ 3	3	3	2/ 3	na(5)	J. Hamalainen
Extensibility						
to no visible artifacts	1	1	1	2	5	J. Fuhrer
to studio-quality data rate	1	1	1	2	5	J. Fuhrer
to higher resolution	2	3	4	2	5	A. Toth
to VHDTV	2	3	4	2	5	A. Toth
to UHDTV	2	3	4	2	5	A. Toth
Provision for future compression enhancement	2	2	2	2	5	J. Bellisio

ATV Interoperability Key: (/ Modified Later)

Evaluation Rating	Proposed Implementation
1	- Easy to implement
2	
3	- Moderately difficult
4	
5	- Very difficult to implement
n	No answer or incomplete
v	Verification needed

IV. Summary of Interoperability Evaluations

Overall

The four all-digital systems were found to be superior to the Narrow-MUSE system with respect to Interoperability and Extensibility on all but two or three characteristics where Narrow-MUSE was approximately equal to the all-digital systems. With respect to Scope of Services and Features, the all-digital systems also ranked better than Narrow-MUSE but the difference was smaller.

Interoperability

AD-HDTV was judged more interoperable with headers/descriptors and packet networks. DSC-HDTV was judged more interoperable with computers with CCDC close behind. CCDC was judged more interoperable with film. The Review Board evaluations for interoperability were averaged with the following results shown to greater precision than justified by the subjective nature of the assessment process:

> 1.88 for AD-HDTV, 1.90 for DSC-HDTV, 2.13 for CCDC, 2.43 for DigiCipher and 3.61 for Narrow-MUSE.

AD-HDTV also led the ranking by the proponents when averaged over all interoperability categories, but DSC-HDTV was ranked by the individual experts best or tied for best in all categories except headers/descriptors.

Scope of Services and Features

AD-HDTV was ranked by the Review Board as best or tied for best in six out of nine features. With uniform weighting, the average evaluations were:

> 1.66 for AD-HDTV, 1.71 for DSC-HDTV, 1.83 for CCDC , 1.85 for DigiCipher and 2.51 for Narrow-MUSE.

It should be noted that the differences are not significant and that uniform weighting may not be appropriate as discussed on page 5. AD-HDTV also received the best average evaluation in the proponents' evaluation. DSC-HDTV was ranked best or tied for best in all features by the individual experts.

Extensibility

DSC-HDTV was ranked by the Review Board as best or tied for best in four out of the six categories and was almost equal to AD-HDTV in a fifth feature. With uniform weighting, the average evaluations were:

> 2.43 for DSC-HDTV, 2.53 for both AD-HDTV and CCDC, 3.08 for DigiCipher, and 4.50 for Narrow-MUSE.

AD-HDTV was ranked by the proponents as best or tied for best in all extensibility categories. AD-HDTV was also ranked best or tied for best by the individual experts.

APPENDIX I

PROPONENT RESPONSES

ATRC SUBMISSION
Glenn Reitmeier

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

	ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
H	**Interoperability** CABLE	1	1	1	1	1
H	with digital technology	1	1	1	1	5
H	with headers/ descriptors	1	3	3	3	5
H	with NTSC	1	3	1	3	4
H	with film	1	1	1	2	3
H	with computers	1	3	4	3	5
H	with satellites	2	2	2	3	5
H	with packet networks	1	3	3	3	5
H	with interactive systems	2	3	3	3	5
H	Format conversion	2	3	3	3	3
M	Scalability	1	4	4	2	5
	Scope of Services and Features					
L	Initial use for ancillary data	1	1	1	1	1
H	Audio	1	1	1	1	2
M	Data	1	3	3	3	5
M	Text	1	3	3	3	5
M	Captioning	1	3	3	3	5
M	Encryption	1	1	1	1	4
H	Addressing	1	3	3	3	4
M	Low-cost receiver	1	3	3	2	1
H	VCR capability	1	3	3	4	2
	Extensibility					
M	to no visible artifacts	2	3	3	3	4
M	to studio-quality data rate	2	3	3	3	3
M	to higher resolution	2	3	3	3	5
M	to VHDTV	2	3	3	3	5
L	to UHDTV	2	3	3	3	5
M	Provision for future compression enhancement	2	3	3	3	5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

ATRC IMPORTANCE WEIGHTING FACTOR: H - HIGH IMPORTANCE
M - MEDIUM "
L - LOW "

NOTES ON INTEROPERABILITY EVALUATION MATRIX

INTEROPERABILITY

CABLE

All proposed systems can be delivered successfully over cable.

DIGITAL TECHNOLOGY

The four digital systems have an extreme advantage here.

HEADER/DESCRIPTORS

AD-HDTV has been designed and tested with header/descriptor capabilities. Header/descriptors could be retrofitted to the other digital systems with some difficulty. Header/descriptors are not meaningful or implementable in an analog system.

NTSC

The 2:1 relationship of the 1050 line systems (AD-HDTV and DigiCipher) to NTSC is an advantage. The 3:2 relationship of the 787 line systems (CCDC and DSC) is not as desirable, but better than the 1125 lines of Narrow Muse.

FILM

AD-HDTV, DigiCipher and CCDC have a film transmission mode at 24 fps. DSC requires conversion to 59.94 fps, but can benefit from some increase in encoding efficiency. Narrow Muse will exhibit traditional TV and film interoperability.

COMPUTERS

AD-HDTV has interoperability at the picture level (in its progressive scan and square pixel mode) and at the compressed bit stream level (MPEG). CCDC and DSC only have progressive scan and square pixel interoperability at the picture level. DigiCipher suffers from interlace and rectangular pixels in all modes, but it at least is digital. As an analog system, Narrow Muse will be very difficult to integrate with computers.

SATELLITE

All digital systems require remodulation (to QPSK) for satellite transmission. AD-HDTV, DigiCipher and CCDC all have constant data rate, while DSC suffers from a variable data rate, making the interfacing more difficult. As an analog system, Narrow Muse will require remodulation to FM, a relatively simple task that is similar in complexity to the QPSK remodulation required in AD-HDTV, DigiCipher and CCDC.

PACKET NETWORKS

Only AD-HDTV has a packetized transmission format that has been implemented and tested. Other digital systems can be interfaced to packet networks with some moderate difficulty. As an analog system, Narrow Muse will be very difficult to integrate with packet networks.

INTERACTIVE SYSTEMS

AD-HDTV has MPEG compression, which will be used in CD-I and other interactive systems. CCDC, DigiCipher and DSC can be interfaced to interactive systems, but may require either multistandard decoders or decompression/recompression (with the associated artifacts). As an analog system, Narrow Muse will be very difficult to integrate with computers.

FORMAT CONVERSION

Picture level format conversion equipment has been in existence for several decades. There is a wealth of knowledge about performance and cost/performance tradeoffs. Therefore, all systems are equal at this level. Conversion among compressed forms is much more difficult, and will likely require decompression/recompression (with the associated artifacts). AD-HDTV has an advantage over the other digital systems because its MPEG compression syntax will be common to other MPEG-based systems, and it will not require decompression/recompression.

SCALABILITY

At the compressed bit stream level, AD-HDTV and DSC both exhibit some scalability, since they have "viewable picture subsets" as part of their two-tier transmission. AD-HDTV carries the scalability to the signal level, since the HP bit stream is a separate carrier (DSC uses a single carrier that is time division multiplexed). CCDC and DigiCipher are inherently single tier systems, although they can have some rudimentary DC term extraction from their DCT.

STUDIO QUALITY

All proposed systems can be extended to provide studio quality versions. AD-HDTV's MPEG compression can accommodate several different approaches within the same compression syntax, depending upon studio requirements (e.g., with a flexible GOP structure, MPEG can have all I-frames to meet editing requirements). MPEG GOP structure also makes studio operations in compressed form easier. Narrow Muse will require encoding/decoding to 1125/60.

HIGHER RESOLUTION

All proposed systems can be extended to provide higher resolution versions. AD-HDTV's MPEG compression is already a standard with this proven capability. As an analog system, Narrow Muse is not well suited to this flexibility to accommodate higher resolution.

VHDTV AND UHDTV

The four digital systems can be extended to provide VHDTV and UHDTV versions. AD-HDTV's MPEG compression is already a standard with the proven capability to increase resolution. VHDTV and UHDTV that is backward compatible with HDTV will require an augmentation approach. AD-HDTV's packet structure provides a mechanism to achieve this, while other digital systems will be more difficult to augment without damaging service to existing receivers. As an analog system, Narrow Muse is not well suited to benefit from analog augmentation approaches.

FUTURE COMPRESSION ENHANCEMENT

The four digital systems can all benefit from future encoder improvements. However, MPEG compression was designed to provide this capability as a design consideration. Narrow Muse is not as well suited to benefit from future encoder improvements.

for MIT/ATVA (MIT has chosen not to score CCDC for possible conflict of int

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	1		1	1	5
with headers/ descriptors	1		2	2	5
with NTSC	2		2	2	2
with film	1		1	1	5
with computers	3		3	3	5
with satellites	2		2	2	5
with packet networks	1		3	3	2
with interactive systems	3		3	3	5
Format conversion	3		3	3	5
Scalability	3		3	2	4
	4		4	4	5
Scope of Services and Features					
Initial use for ancillary data	3		3	3	5
Audio	1		1	1	1
Data	1		1	1	2
Text	1		1	1	2
Captioning	1		1	1	2
Encryption	2		2	2	2
Addressing	2		3	3	3
Low-cost receiver	2		2	2	2
VCR capability	2		2	2	2
Extensibility					
to no visible artifacts	1		1	1	4
to studio-quality data rate	1		1	1	4
to higher resolution	1		2	2	5
to VHDTV	1		2	2	5
to UHDTV	1		2	2	5
Provision for future compression enhancement	3		3	3	4

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

GENERAL INSTRUMENT
VIDEOCIPHER DIVISION

October 6, 1992
3 pages

FAX TO: Guy Beakley
Stellacom, Inc.
Fax: 703/243-1698

CC: Bob Sanderson
Woo Paik
Jerry Heller
Jeff Krauss
Quincy Rodgers

FROM: Bob Rast RMR

SUBJECT: Interoperability Evaluation Matrix

Attached please find the Interoperability Evaluation Matrix, which we have filled in, with our rating of each of the digital proponents.

We have refined the rating scale you have proposed, and wish to share our definition with you. Whereas as your rating scale was based on ease of implementation only, we believe that a fair rating scale should include other factors as well. The set of factors we used are:

- Ease of implementation
  1. Easy
  3. Moderately difficult
  5. Very difficult
- Performance
  1. Clear advantages
  3. OK
  5. Questionable
- How real
  1. Demonstrated
  2. In hardware, not demonstrated or tested
  3. Has proposal, seems OK
  5. Questionable
- Cost effectiveness
  1. Included and cost effective
  3. Not included, probably OK
  5. Looks costly

The following conveys key thoughts affecting the rating of each item:

Interoperability

with digital technology - topic seems redundant against later more specific topics, or ambiguous. All systems digital.

with headers/descriptors - all have proposals, ADTV's is most real, but not fully verified

with NTSC - DigiCipher has a developed, compatible system, is interlaced, and has favorable scan parameters for conversion to NTSC. Through MUSE, Narrow MUSE has demonstrated NTSC compatibility. ADTV has the scanning relationship advantages also.

with film - DigiCipher and CCDC have implemented this feature in prototypes, and demonstrated it . Interlace systems get more benefit from the feature.

with computers - progressive systems have advantage. Square pixels also an advantage, but not as much so. DigiCipher/computer interoperability not as well thought out.

with satellites - DigiCipher has demonstrated satellite transmission, and capability built into CCDC.

with cable - even if cable is being evaluated elsewhere, it is important to any legitimate interoperability assessment. (Note that there are overlaps with other parts of the Advisory Committee in other areas also.) DigiCipher capability has been demonstrated and tested, uses straightforward QAM approach. CCDC is compatible. VSB more desirable than split signal QAM.

with packet networks - nobody has demonstrated, but ADTV most real.

with interactive systems - no particular advantages for anybody.

Format conversion - somewhat easier to convert progressive images.

Scalability - ADTV and DSC have described scalable subset at transwitch level. All scalable at the picture processing level.

Scope of Services and Features

Initial use of ancillary data - no particular advantages.

Audio - Dolby-based systems compatible with Dolby surround sound proposal, MIT audio compatible at interface.

Data - no particular advantages.

Text - no particular advantages.

Captioning - no particular advantages.

Encryption - all amenable, no particular advantages.

Addressing - DigiCipher and CCDC reserved capacity for this.

Low-cost receiver - possible advantage for ADTV and DSC with transwitch layer proper subset, offset by DigiCipher fundamentally lower cost.

VCR capability - DigiCipher has demonstrated, and described tricks in greatest detail (to SS/WP-3). CCDC is compatible with DigiCipher in this. DSC tricks feasibility deserves more detail.

Extensibility

to no visible artifacts - DigiCipher has shown simulations at 30 Mbps, CCDC is related technolgy.

to studio-quality data rate - none have demonstrated, nor done much investigation.

to higher resolution - some advantage to interlace for defined migration to 1050 progressive. MUSE is higher resolution, and Narrow MUSE is derived from it.

to VHDTV - no advantages, some questions.

to UHDTV - no advantages, some questions.

Provision for future compression enhancement - no clear advantages.

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	2	2	2	2	4
with headers/ descriptors	2	3	3	3	5
with NTSC	2	3	1	3	1.5
with film	2	2	1	3	4
with computers	3	2	4	2	5
with satellites	3	2	1	3	3
with packet networks	2	3	3	3	5
with interactive systems	3	3	3	3	3
Format conversion	3	2	3	2	4
Scalability	2	3	3	2	3
Scope of Services and Features					
Initial use for ancillary data	3	3	3	3	3
Audio	3	2.5	2	2	3
Data	3	3	3	3	3.5
Text	3	3	3	3	3.5
Captioning	3	3	3	3	3
Encryption	3	3	3	3	4
Addressing	3	2	2	3	4
Low-cost receiver	2.5	3	2.5	2.5	3
VCR capability	3	2.5	2	3.5	4
Extensibility					
to no visible artifacts	3	2.5	2	3	3
to studio-quality data rate	3	3	3	3	2.5
to higher resolution	2.5	3	2.5	3	2
to VHDTV	4	4	4	4	3.5
to UHDTV	4	4	4	4	3.5
Provision for future compression enhancement	3	3	3	3	3
with Cable	3	1.5	1	2.5	4

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	• Easy to implement
3	• Moderately difficult
5	• Very difficult to implement

from: General Instrument
10/5/92

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse	NOTE
Interoperability						
with digital technology	1	1	1	1	5	(a)
with headers/ descriptors	2	3	3	1	5	(b)
with NTSC	2	1	2	1	3	(c)
with film	2	1	2	1	3	(d)
with computers	3	3	4	1	5	(e)
with satellites	1	1	1	1	3	(f)
with packet networks	1	3	3	1	5	(g)
with interactive systems	3	1	3	1	5	(h)
Format conversion	3	1	3	1	5	(i)
Scalability	3	3	4	2	5	(j)
Scope of Services and Features						
Initial use for ancillary data	1	3	3	1	2	(k)
Audio	1	1	1	1	1	(k)
Data	1	3	3	1	3	(k)
Text	1	2	2	1	2	(k)
Captioning	1	2	2	1	2	(k)
Encryption	1	1	1	1	5	(k)
Addressing	1	1	1	1	3	(k)
Low-cost receiver	4	2	2	2	3	(l)
VCR capability	2	4	4	3	3	(m)
Extensibility						
to no visible artifacts	3	2	3	1	5	(n)
to studio-quality data rate	3	2	3	1	5	(n)
to higher resolution	3	3	4	2	5	(o)
to VHDTV	3	3	4	2	5	(o)
to UHDTV	3	3	4	2	5	(o)
Provision for future compression enhancement	2	3	3	2	5	(p)

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

Zenith/AT&T
9/25/92

NOTES FOR INTEROPERATILITY EVALUATION MATRIX

(a) The four digital systems all share the virtues of digital signal representation, so that the images, parts of images, and data associated with images can be stored, rearranged and reformatted in arbitrary ways depending on the needs of the application. Such flexibility is not available for an analog system.

(b) ADTV and DSC have a larger amount of channel capacity available for use in headers and descriptors. DSC has a larger unassigned channel capacity for ancillary data than the other digital systems, while ADTV appears to have the least such unassigned ancillary data channel capacity.

(c) The conversions from interlaced scans to interlaced scans are more complicated, and are likely to incorporate conversion artifacts. The Narrow-MUSE scanning numbers are not congenial to NTSC conversion.

(d) Film to interlace requires extra 1/2-frame memory for reinterlace at receiver, for ADTV and DigiCipher. Narrow-MUSE has no mechanism for taking advantage of film's lesser frame rate to deliver better pictures.

(e) ADTV and DigiCipher will introduce interlace artifacts. DigiCipher and CCDC do not have a clean packet/slice structure from which to extract data for windowing. Narrow-MUSE uses analog signals.

(f) Digital systems will avoid the addition of noise associated with analog modulation on satellite.

(g) ADTV and DSC have fixed-size packets with priority assignments. DSC has a variable information content rate which is easily carried by packet networks. Therefore, a broadcaster option could adjust the 2-level/4-level mix for applications that take advantage of the variable bit-rates supportable by packet transmission. DigiCipher and CCDC data objects have variable size, no slice boundaries, and have no priority structures.

(h) ADTV and DigiCipher are interlaced, making interactive operation more complicated. ADTV has long round-trip delay. DigiCipher and CCDC incorporate 4 moving panels with progressive refreshing, which makes interactive applications less graceful.

Zenith/AT&T
9/30/92

(i) ADTV is interlaced, so that format conversions are more complicated and introduce artifacts, but has service type indicators. DigiCipher is interlaced. CCDC and DSC are progressive-scanned, facilitating conversions spatially and temporally. DigiCipher motion estimates are less precise (not sub-pixel, not local), making frame interpolation more difficult.

(j) None is scalable in the strictest sense of subsettable data streams over a wide range of performance levels. ADTV and DSC have priority data providing limited scalability. Interlace in ADTV and DigiCipher makes scalability difficult since DCT in an interlaced field will exhibit different artifacts from field to field when only DC coefficients are retained.

(k) The digital systems all can flexibly support these features by appropriate formatting and use of headers and descriptors. The Narrow-MUSE system's audio is digital, but has limited other digital channel capacity for these features, and analog encryption is complicated. CCDC and DigiCipher have only limited capacity and lack supporting data structures for new features.

(l) ADTV requires more memory in receivers, and MPEG decoding with forward and reverse prediction is more complicated. Narrow-MUSE uses a variation of the already-developed MUSE algorithm that is still quite complex, involving analog functions (performed with digital circuits). The other systems have similar processing complexity to each other.

(m) Narrow-MUSE claims digitized version of signal can be recorded using current VCR technology, but high bit rate will be very expensive. ADTV can use I-frames for clean trick modes, although only particular speed (multiples) are supported, and extraction of data may be costly. CCDC and DigiCipher can use PCM data to support fast search, but speed range (intraframe and quality) are limited. DSC can search at any speed but with reduced quality.

(n) Because ADTV and DigiCipher are interlaced, they may require more bits for coding using intraframe only. DSC and CCDC encode progressive-scanned images, are therefore more efficient in moving images. DSC uses vector quantization for coefficient pattern selection, allowing superior matching of coding algorithm to perception in human visual system. Narrow-MUSE extensibility limited by practical analog processing in affordable bandwidths.

Zenith/AT&T
9/30/92

(o) The progressive-scanned CCDC and DSC have better prospects for artifact-free extensions to higher resolution (both spatially and temporally) than the interlaced ADTV and DigiCipher systems. Narrow-MUSE is not flexibly extensive because of the required bandwidth and analog processing. The four panel boundaries, which move from field to field with a total cycle time of 20 fields, and the associated constraints on motion vectors complicate extensibility of CCDC and DigiCipher.

(p) All the digital systems can use a layering approach to incorporate new developments in compression, and all can exploit improvements in motion estimation. The 4-panel processing in CCDC and DigiCipher will make it difficult to extend in a compatible way to higher spatial and temporal resolution. Analog processing limits prospects for Narrow-MUSE.

Zenith/AT&T
9/30/92

APPENDIX II

REVIEW BOARD RESPONSES

Jules Bellisio
10/8/92

1st cut, subject to revision in meeting because several questions are subject to multiple interpretation.

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	2	2	2	2	4
with headers/ descriptors	2	3	3	3	5
with NTSC	1	1	1	1	5
with film	1	1	2	1	3
with computers	1	1	3	1	5
with satellites	1	1	1	1	2
with packet networks	2	3	4	2	N
with interactive systems	2	1	1	1	1
Format conversion	1	1	3	1	5
Scalability	3	4	4	3	3
Scope of Services and Features					
Initial use for ancillary data					
Audio	1	1	1	1	1
Data	1	2	2	2	2
Text	2	2	2	2	3
Captioning	2	2	2	2	3
Encryption	1	1	1	1	3
Addressing	1	1	1	1	1
Low-cost receiver	3	3	3	3	2
VCR capability	2	3	3	3	4
Extensibility					
to no visible artifacts	1	1	2	1	5
to studio-quality data rate	1	3	3	1	5
to higher resolution	2	2	3	2	3
to VHDTV	2	2	3	2	3
to UHDTV	2	2	3	2	3
Provision for future compression enhancement	2	2	2	2	5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

Digital Transmission	1	1	1	1	3
Still Image	1	3	3	1	3
Compressed NTSC	N	N	N	N	N
Baseband Receiver (A&D)	1	1	1	1	5

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	1	1	1	1	3
with headers/ descriptors	1	2	3	2	4
with NTSC	2	1	2	1	3
with film	2	1	1	1	2
with computers	3	2	3	2	4
with satellites	1	1	1	1	1
with packet networks	1	3	3	3	5
with interactive systems	2	2	2	2	3
Format conversion	3	2	2	2	3
Scalability	3	2	3	2	3
Scope of Services and Features					
Initial use for ancillary data	1	1	1	1	1
Audio	1	1	1	1	1
Data	1	2	2	1	2
Text	1	1	1	1	2
Captioning	1	1	1	1	2
Encryption	1	1	1	1	3
Addressing	1	1	1	1	3
Low-cost receiver	3	3	3	3	3
VCR capability	2	3	3	3	3
Extensibility					
to no visible artifacts	3	2	3	2	4
to studio-quality data rate	3	2	3	2	4
to higher resolution	2	2	2	2	3
to VHDTV	3	2	2	2	6
to UHDTV	3	2	2	2	5
Provision for future compression enhancement	2	1	1	2	3

Jules Cohen
10/5/92

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

I believe strongly that market forces are likely to dictate ~~upon~~ interoperability feature availability. Artificial restraints should not be put on production of a least cost receiver. JC

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

Issues Discussion

ATRC ATV Proponent System

	ATV FEATURE	Advanced Digital	M.I.T. CCDC	G.I. Digicipher	Zenith/AT&T DSC	Narrow-Muse
CT	Interoperability with CableTV	2	2	2	2	5
✓GD	with digital technology *	3	2	4	2	5
ML	with headers/descriptors *	3 V	4 V	5 V	3 V	5
AU	with NTSC *	4	1	4	1	5
AT	with film *	2 V	2 V	3 V	3 V	5
ML	with computers *	4 V	2 V	5	2 V	5
AU	with satellites	2	2	2	2	4
JB	with packet networks *	2 V	4 V	5	2 V	5
ML	with interactive systems *	4 V	3 V	5	2 V	5
✓GD	Format conversion *	5	2 V	4	2 V	5
JB	Scalability *	5	3 V	5	3 V	5
	Scope of Services and Features					
JC	Initial use for ancillary data	3 V	3 V	3 V	3 V	5
JC	Audio *	2	1	2	2	4
JC	Data *	2 V	3 V	4 V	3 V	5
✓GD	Text *	5	2 V	5	2 V	5
GH	Captioning	5	2 V	5	2 V	5
CT	Encryption *	2 V	3 V	3 V	2 V	5
CT	Addressing *	2 V	4 V	4 V	2 V	5
GH	Low-cost receiver *	4 V	2 V	4 V	2 V	5
JH	VCR capability *	3 V	3 V	2	2 V	5
	Extensibility					
JF	to no visible artifacts *	3 V	2 V	4 V	2 V	5
JF	to studio-quality data rate *	3 V	2 V	4 V	2 V	5
AT	to higher resolution *	3 V	2 V	4 V	2 V	0
AT	to VHDTV *	3 V	2 V	4 V	2 V	5
AT	to UHDTV *	3 V	2 V	4 V	2 V	5
JB	Provision for future compression enhancement *	3 V	4 V	5	3 V	5

* → SEE ACCOMPANYING DISCUSSION

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

N INCOMPLETE OR NO ANSWER

V VERIFICATION NEEDED

1 ... N ... FOOTNOTES (PARAGRAPHS OF DISCUSSION)

EVERM [illegible]

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability	1	2	1	2	3
with digital technology	1	1	1	1	4
with headers/ descriptors	1	2	2	2	4
with NTSC	1	2	1	2	2
with film	1	1	1	1	2
with computers	2	2	4	2	5
with satellites	1	1	1	1	2
with packet networks	1	3	2	3	5
with interactive systems	2	2	1	2	4
Format conversion	1	1	1	1	2
Scalability	2	5	4	3	
Scope of Services and Features					
Initial use for ancillary data	1	1	1	1	1
Audio	1	1	1	1	1
Data	1	1	1	1	1
Text	1	1	1	1	1
Captioning	1	1	1	1	1
Encryption	1	1	1	1	4
Addressing	1	3	1	3	3
Low-cost receiver	3	5	3	4	4
VCR capability	1	3	3	3	3
Extensibility					
to no visible artifacts	1	1	1	2	5
to studio-quality data rate	1	1	1	2	5
to higher resolution	2	3	3	2	5
to VHDTV	2	3	3	2	5
to UHDTV	2	3	3	2	5
Provision for future compression enhancement	1	2	3	1	5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

Jack Fuhrer 10/6/92

FCC WP4
Review Board Member Interoperability Evaluation

Gary Demos

6 October 1992

Executive Summary

The following discussion is my evaluation of the advanced television (HDTV) proposals before the FCC as part of the WP4 Interoperability review process.

To summarize my comments:

- I asked a number of questions of the proponents, many of these questions were ignored or answered negatively

- The FCC testing process, which is nearing completion, was flawed in that many crucial issues of interoperability were not tested. This is partly due to the fact that the testing process was begun when all the systems were analog. The testing process should be redesigned in light of digital technology and interoperability issues.

The most critical interoperability problems are:

- Two of the digital systems and the one analog system are interlaced, with non square pixels. These systems should be rejected.

- All of the systems are at 59.94 Hz or 60.0 Hz. This is not compatible with computer display interoperability, requiring greater than 70 Hz. No system should be accepted at 59.94 or 60.0 Hz.

- The apparent ability to send movies (24 fps) at higher resolution has not been explored

- The header proposals are not universal, but are buried within the HDTV packet formats

- The packet structures are untested with ATM or other packet networks

- System modularity, which would allow the best of each system, has not been evaluated.

- None of the systems is scalable downward. A lower resolution subset would allow low cost reception at resolution below full quality, but potentially better quality than NTSC.

- Still imagery communication has not been tested, although apparently feasible.

- It could be significant if a partial screen update capability were provided. Some of the systems may have this capability, but it has not been explored or tested. This would allow reduced bit rate presentation of high resolution images during times when other data is being sent (like address authorization lists), or for use by lower data rate devices.

- Compression quality is rapidly advancing. We need only examine the recent two years to get a feel for the rapid pace of development. No proponents offered proposals for handling major advancements in compression technology. In ten years, it is likely that all of the proposed systems will be obsolete in their approach. How will we handle the extensibility issue in order to allow graceful adoption of future advancements in image compression and representation?

• The FCC process has encouraged innovation up to this point by stimulating digital systems to be proposed by the proponents. However, at this point what is required is cooperation and system modification in an orderly manner, with subsequent testing of any such modifications. The FCC process as now constituted IS DESIGNED TO PRECLUDE such cooperation, system modification, and subsequent testing! Thus, the work of this interoperability review board is likely to have little significance unless the FCC process is amended to take interoperability seriously.

Summary Of My Evaluation Of The Proponents

My evaluation is presented in much more detail below. However, a summary of my view of each system is presented here:

	Strengths	*Weaknesses*
Accept If Modified:		
AT&T/Zenith	Progressive Scan Square Pixels Ability To Correct ATM Prioritized Data 8 x 8 small motion vector blocks Sub-Pixel motion resolution Vector Quantization Easy NTSC DownConversion Digital	59.94 Hz No 24 Hz mode Insufficient ATM Correction Odd Format Size (523/262) Regional leak not fully tested Buried Header Not Universal Partial Image Update Not Tested Dependent Audio Channels Not Scalable
MIT/ATVA	Progressive Scan Square Pixels 24 Hz Film Mode Independent Audio Channels Sub Pixel motion resolution Both 8 x 8 and 16 x 16 motion Easy NTSC DownConversion Digital	59.94 Hz No ATM Mapping Data Not Prioritized Odd Format Size (525 Datalines) No Header Proposed Partial Image Update Not Tested Sliding Panels Not Scalable
Reject As Not Interoperable:		
ATRC	Packet Structure Prioritized Data Digital	Interlaced Scan Non-Square Pixels Provision For Square Confused 59.94 Hz Header Under ATRC Packets Dependent Audio 16 x 16 coarse motion only Not Scalable
G.I.	Digital	Interlaced Scan Non-Square Pixels 59.94 Hz No Header Dependent Audio 16 x 32 coarse motion only Sliding Panels Not Scalable
NHK		Analog Interlaced 59.94 Hz

No Header
No Data Formats
Not Scalable

Unanswered Questions

• I submitted a list of questions which I was concerned with to the proponents prior to the review meeting. A copy of my questions are attached to this evaluation. My questions were answered in writing and by presentation by MIT. None of the other proponents answered the questions directly, although some of the questions were answered in the course of presentations.

• In general, the answers which I was able to get fell into the following categories:

1) Seems reasonable, have potential solution, but haven't tried it

a) Could be tested without undue difficulty

b) Would be difficult to test with existing prototype equipment

2) Good issue, haven't considered it.

3) Not able to support the item concerned.

• Some of the proposals for solutions seem too preliminary to be plausible

I will cover these issues in more detail here.

Testing Process

• There have been substantial flaws in the testing process from an interoperability point of view. These are:

• The ATTC only provided 59.94 Hz signals to the proponents. No other rates were explored.

• 24 Hz signals were not supplied even though some of the systems can accept and transmit at 24 Hz

• Increased resolution at 24 Hz vs higher rates has been proposed, but has not been tested.

• 70+ Hz computer display compatibility was not tested in any way

• Still frames were not tested. Quality, how many per second, etc. should be tested.

• No flexibility in performance was tested. Such flexibility would include:

• exploring the limits of pixels per second (Jae Lim's comments were that he could go as high as 100 MPixels/second).

• exploring conditional replenishment

• exploring alternate frame rates

• Potential modularity of systems was not tested. Systems were tested only as a whole.

• Header mechanism has not been tested for reliability (error performance), universality, usefulness.

• ATM mapping and error performance was not tested.

• Computer text and graphics on screen was not tested (PC or computer workstation display). Overlay plane implementations (if used) were not implemented or tested.

• Scalability not proposed and not tested.

• Extensibility not tested. Only proposed verbally.

• NTSC down conversion not tested for quality.

• "Production" or "Contribution" quality extensible superset not tested, although proposed verbally.

• The "high priority subsets" used in ATRC and AT&T/Zenith were not tested for viewability by themselves, although private demonstrations of viewability have been promised.

• The distribution of an address list for pay-per-view cable use would require a lower bit rate to be used for still image or partial screen update during the distribution. The distribution of a million subscriber address list should be tested, and the reliability of delivery should be measured in terms of the number of subscriber errors. The quality of the image during the distribution should be tested.

• Encryption may affect the error performance. The impact of the use of encryption on the quality of the picture should be tested. The protection afforded by proposed encryption schemes should be evaluated. The interaction of headers and encryption should be tested, including "in the clear" headers in an otherwise encrypted data stream for such uses as authorization of decryption to new users being added on.

• A key use of advanced television may be professional consultation and collaborative work. A multi person teleconference should be tested. An example might be a medical consultation.

• Some of the system proponents indicated compatibility with digital compressed NTSC proposals to Cable Labs. Such compatibility should be tested for receiver compatibility, distribution compatibility, and interoperability compatibility. Some claimed that the compressed NTSC standard would only be selected after the advanced television standard. This should be verified if true, but testing for interoperability and compatibility will still need to occur prior to selection of any given compression algorithm.

• The use of a VCR, laser disk, or other devices implies a digital port, wire, and signal. The specifications, reliability, and efficiency of such connections should be evaluated prior to selection of any such signal designations. If industry standard digital formats are used (such as IEEE P1394), then these signals should be tested on these formats.

• There was substantial discussion of higher quality signal feeds. Channels which might have wider bandwidth than 20 Mbits per second include ATM networks, long haul fiber (e.g. SONET), cable TV, satellite transponder channels, cellular digital, OFDM, or other networks. The proposals should be tested with these other data rates for compatibility and quality. The mapping of reduced data rate extractions of 20 Mbits/sec from higher data rates should be tested. The methods for simultaneously generating 20 Mbits/sec in addition to higher rates from a source signal should be tested.

• Some people, such as Jim Clark, Chairman of Silicon Graphics, feel that 3-D graphics will be routinely affordable for home receivers and computers. No testing of 3-D graphics interoperability was investigated.

PANASONIC TECHNOLOGIES Inc.
Matsushita Applied Research Laboratory

95D Connecticut Drive, Burlington, NJ 08016 USA

Telephone 609-386 5995 Fax 609-386 4999

TO: MR. G. BEAKLY
STELLA COM

FROM: JUKKA HAMALAINEN
MARL

SUBJ: PS-WP/4 INTEROPERABILITY REVIEW

PAGES: 3

Attached you will find my ATV Interoperability Evaluation Matrix. In addition to the Matrix you will find below, some of my thoughts and my guidelines used for completing the Matrix.

1. First of all, the present deadline for the US HDTV standard is close. Therefore, the proponents cannot make any major hardware nor software changes.

2. Changes which effect the tested video and audio performances cannot be accepted without having the proposals at least partly retested.

3. The cost ratio between the decoder and the rest of the receiver (display, power supply, cabinet, audio amplifiers, loudspeakers - RF tuner? - etc.) will be about 30% - 70%, perhaps even more in the later phase. Therefore, we can accept added complexity in the system if this improves the scalability and the extensibility of the system. In the above example, increasing the cost of the decoder by 50% would increase the cost of the receiver 15%.

4. The complexity of the system cannot however go so far, that the cost of the encoder gets out of hand or makes a portable encoder impractical in the future.

5. The selected systems should not favor NTSC too much. Eventually, NTSC will be disbanded. The artifacts converting from HDTV to NTSC will be masked by the NTSC systems own artifacts. The basic frame rate of 60Hz (or even 72 Hz with a hierarchy of 12, 24, 48, 72) could enhance US position as the main provider of TV programs (movies and live TV as well) in the world.

6. Square pixels, this issue which has caused many heated discussions, should be discussed once more in light of interoperability. My dilemma is as follows:
   - The present systems have been tested using analog signals as an input to the compression encoder, each proponent then digitizes this video signal providing the scanned image.

10/05/1992 16:11 MARL BURLINGTON NJ USA 609 386 4999 P.01

- In the near future all studio processing will be digital, and the video as well as the other data coming to the encoder will be digital.
- The production standard for the studios, should it be the same as the transmission standard (active pixels per frame)? Traditionally in the analog world the production quality was higher than the transmission quality. To maintain square pixels for both transmission and production can only be done if the numbers stay the same, if the production quality has to be better, then both the line and pixel rates have to be modified to maintain the square pixels.

7. Packetizing the data is essential for good network operations as well as for VCR's and Computers.

8. Regarding (receiver) scalability, I would like to classify the receivers, and video quality as follows:

• <10 inch,	portable receivers
• 20 - 30 inch	regular viewing
• >30 inch	HiFi receivers

I don't think that DC component of DCT is adequate for any viewing.

9. Can the encoder provide outputs for:
   a) 200-400 Mbs "lossless compression," without temporal compression enabling simple effects such as wipes and insertions.
   b) 40-45 Mbs distribution level, allowing cuts at edit points.

10. Almost everyone agrees, that eventually the systems will be progressively scanned. Either this has to be put into the systems from the beginning or the proponents have to show how the older receivers can handle this type of data. Interlace will provide artifacts on the computer displays.

11. I hope that the promises and proposals given for the added interoperability do not change too much the numbers established by WP-3 for the cost analysis.

12. A few words about the systems.
    - Because the Narrow Muse is an analog system, it cannot be rated for digital interfaces.
    - Digicipher and CCDC four have vertical panels, this provides nice commonality for components and works well with their plans for NTSC transmissions, reduces cost, but what about the future expandability?

Finally, I personally favor a computer compatible approach, but within reason, TV for most consumers will be a one way media and not a LAN type network, although it is getting much more sophisticated, pay per view type operations and certain other low level interactive functions will become more common. The ratio between regular viewing and computer type operations will remain high (perhaps 100:1 or more), therefore we cannot penalize the majority of viewers by increasing the cost too much or by making the viewing too complicated.

Sincerely,

Jukka

Jukka Hamalainen

JKH/jay
cc: Bob Sanderson

10/05/1992 16:12 MARL BURLINGTON NJ USA 609 386 4999 P.02

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability w. cable	1	1	1	1	1
with digital technology	1	1	1	1	NA (5)
with headers/ descriptors	1	3	3	1	NA (5)
with NTSC	2	1	1	2	3
with film	2	1	1	2	3
with computers	3	3	3	2	3
with satellites	1	1	1	1	1
with packet networks	1	4	3	2	NA (5)
with interactive systems	2	2	2	2	4
Format conversion	2	2	3	2	5
Scalability	2	4	4	2	5
Scope of Services and Features					
Initial use for ancillary data	These are basic applications for all systems = 1				
Audio					
Data					
Text					
Captioning					
Encryption					
Addressing					
Low-cost receiver	2	3	3	2	2
VCR capability	2	3	3	2	NA (5)
Extensibility					
to no visible artifacts *	3	3	3	3	5
to studio-quality data * rate	3	3	3	3	5
to higher resolution **	4	3	4	3	5
to VHDTV **	4	3	4	3	5
to UHDTV **	4	3	4	3	5
Provision for future compression enhancement	3	3	3	3	5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

* very subjective issues

** preference for progressive systems

J. Hamalainen
Oct 5, 1992

G. Hanover

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse	note
Interoperability cable	2	1	1	1	1	
with digital technology	1	1	1	1	4	a
with headers/ descriptors	1	3	3	3	3	b
with NTSC	3 ✓	3 ✓	3 ✓	3 ✓	3 ✓	c
with film	3	1	1	2	3	d
with computers	3	1	3	1	3	e
with satellites	1	1	1	1	1	
with packet networks	1	3	3	3	5	
with interactive systems	4	4	4	4	4	f
Format conversion	3 ✓	4 ✓	4 ✓	3 ✓	4 ✓	g
Scalability	3	4	4	3	2	h
Scope of Services and Features						
Initial use for ancillary data	1	1	1	1	1	
Audio						
Data						
Text						
Captioning						
Encryption					4	
Addressing					2	
Low-cost receiver	3	3	3	2	2	i
VCR capability	2 ✓	2 ✓	2 ✓	2 ✓	2 ✓	
Extensibility	N ✓					
to no visible artifacts						
to studio-quality data rate						
to higher resolution						
to VHDTV						
to UHDTV						
Provision for future compression enhancement						
Comp NTSC	NV	NV	NV	NV	NV	

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Digital xmiss.	1	1	1	1	4
rcvr baseband (A+D)	3 ✓	2 ✓	2 ✓	2 ✓	3 ✓
still image	N	N	N	N	N

a) Spectrum efficiency of AD-HDTV suffers on cable because of co-channel notch.

b) Considerations were: easy reassignment of bits, easy addition of other services and general format of information.

c) All need conversion. NTSC information is not immediately accessible from compressed data stream. Conversion to NTSC is most important to consumer.

d) AD-HDTV requires scan change in receiver. DSC ~~has~~ doesnot have a film mode~~s~~, but can recognize film in its basic algorithm.

e) Considerations were: square pixels and progressive scan.

f) 100 ms is needed for channel change (and preferably other) interactive functions.

g) MPEG and 240M were major considerations.

h) Access to program information to make simple receiver so that receiver need not parse.

i) Considerations

1) System functions (i.e. cost) should be in other parts of the system as much as possible so that all receivers can be low cost.

2) Scalability and especially extensibility and interoperability must not eliminate the possibility of plain vanilla TV sets.

3) Concept of zero cost default - After high definition requirements have been satisfied in the system, other useful information is formatted in such a way that there is not additional cost to products which do not use that information.

4) Same as note h

5) Most cost of receivers are not in signal processing and the cost of this section is the easiest to cost reduce. The proposed system must have little impact on other parts of the receiver than the DSP section(s).

Apple Computer, Inc.
Advanced Technology Group

Mike Liebhold
Media Architecture Research

ACATS PS-WP4 Interoperability Review Notes

October 6, 1992

Interoperability with 'Computers'

'Digital technology' and *interactive systems'* as used in the matrix are ambiguious. No scores are given. Answers given for *'computer' interoperability '* should apply

Definition of 'Computer'

The base assumption herein is that functional differences are rapidly diminishing between *computers* , television and telephones. The term *'computer '* used here refers arbitrarily to *intelligent* devices for display, creation, and communication of mixed data and media. Computers are handheld, on desktops, TVtops, and eventually, displayed on flat wall screens.

Basis for Ratings

ATV systems' interoperability *with computers* is determined here by averaging separate factors determining complexity and economy of transcoding between applications, media and channels:

- Syntax for headers/descriptors and other data structures
- Use of square sampling grids (square pixels)
- Progressive or interlace image display
- Spatial Resolution
- Display refresh rate
- Scalability
- Interoperability with packet networks

Worksheets for each of the proponents are attached

Advanced Digital HDTV
Digital Spectrum-Compatible HDTV
Channel-Compatible Digicypher HDTV
Digicypher HDTV
Narrow-MUSE HDTV

Mike Liebhold -
Apple Computer, Inc

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV FEATURE	ATV Proponent System: Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability w. CATV	1	1	1	1	1
with digital technology	*	*	*	*	*
with headers/ descriptors	2	3N	3N	3N	5N
with NTSC	2	2	2	2	4
with film	2V	1	3	3	5
* with computers	2.42V	2.42NV	3.1NV	2.28NV	3.7V
with satellites	1	1	1	1	3
with packet networks	2V	3N	3N	2V	5
with interactive systems	*	*	*	*	*
Format conversion	3	3	4	3	5
Scalability	3V	3V	4V	3V	4V
Scope of Services and Features					
Initial use for ancillary data	3	3	3	3	5
Audio	2	2	2	2	2
Data	3	3	3	3	5
Text	-	-	-	-	-
Captioning	2	2	2	2	2
Encryption	2V	1V	1V	2V	2
Addressing	2V	1V	1V	2V	3
Low-cost receiver	3V	3V	4V	3V	4V
VCR capability	2	2	2	2	2
Extensibility					
to no visible artifacts	3	2	3	2	4
to studio-quality data rate	3	2	3	2	3
to higher resolution	4	4	4	4	5
to VHDTV	4	4	4	4	5
to UHDTV	4	4	4	4	5
Provision for future compression enhancement	4	4	4	4	5

N,V

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

N - incomplete
V - verification needed
* - see attached

Advanced Digital HDTV*AD-HDTV*

Score	
2.42V	• Average
	• Syntax for headers/descriptors and other data structures
2	MPEG syntax is well defined, but not universal
	• Use of square sampling grids (square pixels)
2V	non-square (square mode needs verification)
	• Progressive or interlace image display
2	interlace (supports progressive at lower data rate)
	• Spatial Resolution
3	1440x960 not easily transcoded
	• Display refresh rate
3	59.94 not easily transcoded
	• Scalability
3V	value of HP stream to be determined
	• Interoperability with packet networks
2V	needs verification

Channel-Compatible Digicypher HDTV *CCDC-HDTV*

<u>Score</u>

2 .42NV	• <u>Average</u>
	• Syntax for headers/descriptors and other data structures
3N	syntax is not defined, may be improved (?)
	• Use of square sampling grids (square pixels)
1	square
	• Progressive or interlace image display
1	progressive
	• Spatial Resolution
3	1280x720 not easily transcoded
	• Display refresh rate
3	59.94 not easily transcoded
	• Scalability
3V	value of DCT to be determined
	• Interoperability with packet networks
3N	needs additional data

Digicypher HDTV *DC-HDTV*

Score	
3.1NV	• Average
	• Syntax for headers/descriptors and other data structures
3N	Syntax is not well defined, but may be improved (?)
	• Use of square sampling grids (square pixels)
3	non-square
	• Progressive or interlace image display
3	interlace (verification required of alternate display)
	• Spatial Resolution
3	1440x960 not easily transcoded
	• Display refresh rate
3	59.94 not easily transcoded (verification required of alternate frame-rate)
	• Scalability
4V	to be determined
	• Interoperability with packet networks
3N	Additional data needed

Digital Spectrum-Compatible HDTV *DSC-HDTV*

Score

2.28NV	• Average	
	• Syntax for headers/descriptors and other data structures	
3N		syntax is not well defined, may be improved (?)
	• Use of square sampling grids (square pixels)	
1		square
	• Progressive or interlace image display	
1		progressive
	• Spatial Resolution	
3		1280x720 not easily transcoded
	• Display refresh rate	
3		59.94 not easily transcoded
	• Scalability	
3V		value of scalable DCT to be determmined
	• Interoperability with packet networks	
2N		additional data needed

Narrow-MUSE HDTV *Narrow-Muse*

Score	
3.7V	• Average
	• Syntax for headers/descriptors and other data structures
5	Syntax is not defined
	• Use of square sampling grids (square pixels)
3	non-square
	• Progressive or interlace image display
3	interlace
	• Spatial Resolution
3	1920x1035 not easily transcoded
	• Display refresh rate
3	59.94 not easily transcoded
	• Scalability
4V	value of sub-sampled four-field to be determined
	• Interoperability with packet networks
5	not suitable

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV FEATURE	ATV Proponent System: Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability w/ Cable	1	1	1	1	1
with digital technology	1	1	1	1	5
with headers/ descriptors	1	1	1	1	5
with NTSC	1	1	1	1	1
with film	1	1	1	1	3
with computers	3	1	1	1	5
with satellites	1	1	3	1	1
with packet networks	1	3	3	3	5
with interactive systems	3	3	3	3	3
Format conversion	3	3	3	3	3
Scalability	5	5	5	5	5
Scope of Services and Features					
Initial use for ancillary data	N	N	N	N	N
Audio	1	1	1	1	3
Data	1	1	1	1	1
Text	1	1	1	1	1
Captioning	1	1	1	1	1
Encryption	1	1	1	1	3
Addressing	1	1	1	1	3
Low-cost receiver	5	5	5	5	5
VCR capability	3	3	3	3	5
Extensibility					
to no visible artifacts	1	1	1	1	3
to studio-quality data rate	1	1	1	1	3
to higher resolution	5	5	5	5	5
to VHDTV	5	5	5	5	5
to UHDTV	5	5	5	5	5
Provision for future compression enhancement	1	1	1	1	5

N - Incomplete or No Answer
V - Verification Needed

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

Craig K. Tanner
Cable Television Laboratories

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

	ATV Proponent System				
ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability with CATV	1	1	1	1	1
with digital technology	1	1	1	1	3
with headers/ descriptors	1	3	3	1	5
with NTSC	2	2	2	2	3
with film	1	2	2	1	4
with computers	2	2	5	2	5
with satellites	1	1	1	1	1
with packet networks	1	3	3	2	5
with interactive systems	1	2	2	2	5
Format conversion	2	3	3	3	5
Scalability	4	5	5	5	5
Scope of Services and Features					
Initial use for ancillary data	1	1	1	1	1
Audio	1	1	1	1	1
Data	1	1	1	1	1
Text	1	1	1	1	1
Captioning	1	1	1	1	1
Encryption	1	1	1	1	1
Addressing	1	1	1	1	1
Low-cost receiver	3	2	2	2	3
VCR capability	3	3	3	3	3
Extensibility					
to no visible artifacts	2	2	2	2	5
to studio-quality data rate	2	2	2	2	5
to higher resolution	2	3	4	2	5
to VHDTV	2	3	4	2	5
to UHDTV	2	3	4	2	5
Provision for future compression enhancement	3	3	3	3	4

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

A. Toth
Eastman Kodak

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability CABLE	1	1	1	1	1
with digital technology	1	1	1	1	5
with headers/ descriptors	1	3	3	1	5
with NTSC	2	1	2	1	3
with film	1	1	1	1	5
with computers	3	3	3	3	5
with satellites	2	2	2	2	2
with packet networks	1	3	3	1	5
with interactive systems	?	?	?	?	?
Format conversion	2	3	3	2	4
Scalability	4	4	4	4	4
Scope of Services and Features					
Initial use for ancillary data	3	3	3	3	5
Audio	1	1	1	1	1
Data	1	1	1	1	3
Text	1	1	1	1	3
Captioning	1	1	1	1	3
Encryption	3	3	3	3	3
Addressing	1	1	1	1	3
Low-cost receiver	2	3	3	3	2
VCR capability	2	3	3	3	3
Extensibility					
to no visible artifacts	2	2	2	2	4
to studio-quality data rate	2	2	2	2	4
to higher resolution	3	3	3	3	5
to VHDTV	3	3	3	3	5
to UHDTV	3	3	3	3	5
Provision for future compression enhancement	3	3	3	3	5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	• Easy to implement
3	• Moderately difficult
5	• Very difficult to implement

CAPITAL CITIES / ABC., INC. OCT 5, 1992

ATV INTEROPERABILITY EVALUATION MATRIX

Introduction:

Of the ten performance evaluation criteria, which the FCC will be using to reach its decision on which of the five proposed ATV system could best serve the terrestrial broadcast industry, three have recently been proposed by other potential users of the electronic media: *Interoperability, Extensibility* and *Scope of Services*.

It is being assumed that digital advanced television is more "flexible" than today's analog television and that it could offer the capability of providing useful and *cost-effective* interchange of electronic image, audio and associated data among different signal *formats*, among different transmission *media*, among different *applications*, among different *industries*, among different performance levels, and would therefore meet most of the requirements of *Interoperability*.

It is also assumed that the digital representation of high-resolution images would have the properties that permit future improvements in *performance* or *format* while retaining some measure of interoperability and therefore would comply with the concepts of *extensibility*. The digital representation of the images should preferably be of a hierarchical nature of spatial resolution, temporal rates and image aspect ratios so that a desirable subset of the coded image data stream can be selected for transmission, storage or display as needed for a specific application. This then could meet the objectives of *scalability*.

The third requirement, *scope of services* is related to the channel capacity for ancillary data supported by the ATV system.

Presentations by Proponents:

All presentations followed the ATV Interoperability Evaluation Matrix and attempted to explain how well the ATV systems met or could meet the demanding requirements listed in this matrix. Not addressed were the complexity issues of reformatting the data and the cost implications and associated performance compromises which the missing features/requirements would impose if they were implemented.

It is necessary to prioritize this matrix as most ATV receivers will be used for broadcast/cable reception. For these applications receivers should not be burdened with extra circuitry that may never be used. However, the image data stream could be structured to allow the most important features of the matrix to be implemented if it does not compromise the performance and cost-effectiveness of the primary application: ATV terrestrial broadcast reception. The cost to incorporate these additional features, not relevant to terrestrial broadcasting, should be clearly stated.

It appears that the main goal of all digital system proponents was to design, in a very limited time, a prototype ATV system that would be robust enough for digital terrestrial transmission in a 6 MHz RF channel. To make this possible, very high video compression ratios are used and this may make less efficient scalable compression algorithms less likely to be implemented. From the discussions, it appeared that true *scalability* - different bit rates and different quality levels - is not present in any of the systems.

Although not part of this evaluation, broadcasting requires a set of related compression levels for studio applications: recording, film processing, post-processing, satellite distribution to local affiliated stations at a bit rate high enough to allow further processing. Suggested data rates for these applications could be between 200 Mbs and 45 Mbs. These requirements could impact the choice of compression algorithm for terrestrial transmission.

At such time that a clearly defined list of required features is agreed upon, proponents should be given additional time to verify if and at what cost these features can or should be implemented. It should be noted that some of the features listed in the evaluation matrix were not even known and others not defined when the ATV systems were well advanced in their development to meet the ACATS Lab test deadlines. Proponents may also decide not to implement certain features of the matrix. This too should be made known.

Since the ATV receiver will be multi-purpose; ATV, NTSC, VCR, cable friendly, it is important that for ease and cost effective conversion between ATV and NTSC in both directions the scanning parameters should have a simple relationship to each other so that television signals of different format and sources can be processed for use by a common display. Also, the signal format should be such as to make cost effective VCRs possible and these VCRs should provide the same features as present day VCRs.

Progressive scanning and square pixels offer some powerful advantages over interlace scanning for all forms of signal processing but the trade-off may be *extensibility*. Future upgrade from interlace to progressive scanning using the DigiCipher or ADTV parameters may be easier to accomplish than upgrading to an as yet unknown scanning format using the Zenith/AT&T or MIT parameters.

Capital Cities/ABC, Inc.
October 5, 1992

Narrow Muse

ATTACHMENT IX

Edward Horowitz

ATV Interoperability Evaluation Matrix

Issues Discussion

ATV Proponent System

	ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
CT	Interoperability with Cable TV	2				3
GD	with digital technology	6 depm				2
ML	with headers/ descriptors					3/4 in data channel
AU	with NTSC	2				4/5
AT	with film	4				4/5
ML	with computers				268M	2/3
AU	with satellites	2				2/3
JB	with packet networks	/				Not Applicable
ML	with interactive systems					3
GD	Format conversion					4
JB	Scalability				family	2
	Scope of Services and Features					
JC	Initial use for ancillary data					
JC	Audio					1
JC	Data					✓
GD	Text					1
GH	Captioning					2
CT	Encryption	2				1
CT	Addressing	2 10 million				3
GH	Low-cost receiver				Reduced quality	3
JH	VCR capability					2
	Extensibility					—
JF	to no visible artifacts					2
JF	to studio-quality data rate				spectrum inefficient	3/4
AT	to higher resolution			muse	K to T	4
AT	to VHDTV					digital only
AT	to UHDTV					"
JB	Provision for future compression enhancement				Not particularly easy	4/5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	• Easy to implement
3	• Moderately difficult
5	• Very difficult to implement cost effectively

Over

VOD - possible - "3" level of difficulty

H 261 - likely ~~#~~ able to be done but not Absolute

Conversion - Technically being done today - However to NTSC it is not elegant or cost effective

Note: Because of the existance of a lot of equipment
ie: Film/Telecine
Muse E, T, N, 4 ch. decoders
HDTV to NTSC converters

Technical feasibility is quite often demonstratable
However there is a great deal of uncertainty about cost and mass production.

Edward Horowitz - SARNOFF

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	1				
with headers/ descriptors	1				
with NTSC	1				
with film	1				
with computers	MPEG 1				
with satellites	1				
with packet networks	1				
with interactive systems	1				
Format conversion	2				
Scalability	1				
Scope of Services and Features					
Initial use for ancillary data					
Audio	1 256				
Data	1 } 256				
Text	1 }				
Captioning	1 (9.6)				
Encryption	1				
Addressing	1 (rate?)				
Low-cost receiver	1				
VCR capability	2 (pix/rdp?)				
Extensibility					
to no visible artifacts	1				
to studio-quality data rate	1				
to higher resolution	1				
to VHDTV	2				
to UHDTV	2				
Provision for future compression enhancement	2				

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

How many receivers could be added / sec.

zn

"N" - No Answer
V = Verification needed

E Horowitz
MIT

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	2	2			
with headers/ descriptors	~~3~~ SMPTE	3 SMPTE			
with NTSC		1			
with film		1			
with computers		3			
with satellites		2			
with packet networks		Not known			
with interactive systems		Not known			
Format conversion		5			
Scalability		3			
Scope of Services and Features					
Initial use for ancillary data		2			
Audio		1			
Data		2			
Text		3			
Captioning		2			
Encryption		2			
Addressing		3-4			
Low-cost receiver		No experience			
VCR capability		2			
Extensibility					
to no visible artifacts		1			
to studio-quality data rate		2			
to higher resolution		3			
to VHDTV		3			
to UHDTV		3			
Provision for future compression enhancement		2			

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

(over)

Issues:

(A) Question depth of competence of development Team in translating an ACAdemic approach to actual system. Where is 6.7 contribution?

(B) Modular approach is good source coding is independent of transport/transmission layer.

(C) ~~Plate~~ External box

(D) No Graceful degradation.

APPENDIX III

REVISED REVIEW BOARD RESPONSES

PANASONIC TECHNOLOGIES Inc.

Matsushita Applied Research Laboratory

95D Connecticut Drive, Burlington, NJ 08016 USA

Telephone 609-386 5995 Fax 609-386 4999

TO: MR. GUY BEAKLEY

FROM: JUKKA HAMALAINEN
MARL

SUBJ:

PAGES: 1

Dear Guy,

Attached you will find my updated chart. Not too many changes!

The more we have studied the systems, the closer they get - the same basic principles apply to all to a large extent, DCT, etc.

This became very obvious when my "sister" lab ATVL (mainly for receiver research) made a cost comparison for compression encoders for WP-3.

In addition to the attached chart, I have the following comments:

1. General extensibility
   - GI has limited color resolution.
2. Computer rating
   1. DSC
   2. ADHDTV
   3. GI
   4. CCDC
3. General flexibility
   1. DSC
   2. ADHDTV, the MPEG1 frame structure might hinder future developments.
   3. GI too much emphasis on NTSC
   4. CCDC only considers TV transmission.

Finally I have attached here two charts, which I have used for my general discussions for comparing the basic (video) features of the proposed systems.

Sincerely,

Jukka

Jukka Hamalainen

JKH/jay

November 2, 1992

11/02/1992 10:28 MARL BURLINGTON NJ USA 609 386 4999 P.02

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV Proponent System

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability					
with digital technology	1	1	1	1	1
with headers/ descriptors	1	1	1	1	NA
with NTSC	1	1	1	1	1
with film	2	1	1	2	2
with computers	2	1	2	1	2
with satellites	1	1	1	1	1
with packet networks	1	2	2	2	4
with interactive systems	2	2	2	2	3
Format conversion	2	2	2	2	1~5
Scalability	2	2	2	2	3
Scope of Services and Features					
Initial use for ancillary data	1	1	1	1	1
Audio	1	1	1	1	1
Data	1	1	1	1	1
Text	1	1	1	1	1
Captioning	1	1	1	1	1
Encryption	1	1	1	1	1
Addressing	1	1	1	1	1
Low-cost receiver	2	3	3	2	1
VCR capability	2	2	2	2	1
Extensibility					
to no visible artifacts	2	2	2	2	2
to studio-quality data rate	2	2	2	2	3
to higher resolution	2	2	2	2	2
to VHDTV	4	4	4	4	5
to UHDTV	4	4	4	4	5
Provision for future compression enhancement	2	2	2	2	3

Keiichi Kubota (NHK)

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

ATTACHMENT IX

ATV Interoperability Evaluation Matrix

ATV FEATURE	Advanced Digital	CCDC	Digicipher	DSC	Narrow-Muse
Interoperability w. cable	1	1	1	1	1
with digital technology	1	1	1	1	NA (5)
with headers/ descriptors	1	3	3	1	NA (5)
with NTSC	2	1	1	2	3
with film	2	1	1	2	3
with computers	3	3	3	2	3
with satellites	1	1	1	1	1
with packet networks	1	4	3	2	NA (5)
with interactive systems	2	2	2	2	4
Format conversion	2	~~2~~ 3	3	2	5
Scalability	2	4	4	2	5
Scope of Services and Features					
Initial use for ancillary data					
Audio					
Data					
Text	These are basic applications for				
Captioning	all systems = 1				
Encryption					
Addressing					
Low-cost receiver	~~2~~ 3	3	3	~~2~~ 3	~~2~~ 3
VCR capability	~~2~~ 3	3	3	~~2~~ 3	NA (5)
Extensibility					
to no visible artifacts*	3	3	3	3	5
to studio-quality data* rate	3	3	3	3	5
to higher resolution **	4	3	4	3	5
to VHDTV **	4	3	4	3	5
to UHDTV **	4	3	4	3	5
Provision for future compression enhancement	3	3	3	3	5

ATV Interoperability Key:

Evaluation Rating	Proposed Implementation
1	- Easy to implement
3	- Moderately difficult
5	- Very difficult to implement

* very subjective issues

** preference for progressive systems

J. Hamalainen
~~Oct 5 1992~~
Oct 31/92

11/02/1992 10:28 MARL BURLINGTON NJ USA 609 386 4999 P.03

SOME BASIC PARAMETERS OF THE DIFFERENT TV STANDARDS.

	NTSC	CCIR 601	Widescreen 18 MHz	Narrow MUSE	HDTV CCDC/ DSC	HDTV Digicipher	HDTV AD-HDTV
	Interlaced	Interlaced	Interlaced	Interlaced	Progressive 59.94Hz	Interlaced	Interlace (sq pixels)
Y Pixels/ Active Line	440[1]	720	960	1188	1280	1408	1440
Active Lines	485	485	485	650[2]	720	960	960 (810)[3]
Y Pixels/ Frame	213,400	349,200	465,600	772,200	921,600	1,351,680	1,382,400 (1,166,400)
Net video data rate Y+C[4]	NA	167 Mb/s (8bits)	280 Mb/s (8bits)	Not avail.	662.9/745.7 Mb/s	405.1 Mb/s	517.9 Mb/s (Y=9bits)
Video Bandwidth	4.2 MHz	5.5 MHz	7.5 MHz	20.0 MHz (est.)	34.0 MHz	21.5 MHz	22.0Mhz

Notes: 1) Based on analog luminance bandwidth of 4.2 MHz
2) Based on 650 TV lines processed in the MUSE encoder
3) Proposed for square pixels, 24 and 29.94 fps progressive scanning
4) Numbers based on bit rates during active TV lines only

MARL

Widescreen 18 MHz and D1 versus HDTV Standards

	Line length	Video bandwidth Nyquist*	Video bandwidth filtered**	Pixels per active line sampled	Pixels per active line filtered	Pixels per Frame
Interlaced systems						
AD-HDTV	26.7 usec	27.0 MHz	22.0 MHz	1440	1173	1.38 M
GI	26.7 usec	26.4 MHz	21.5 MHz	1408	1147	1.35 M
Widescreen 18.0 MHz						
Converted to 1050	26.7 usec	18.0 MHZ	15.0 MHz	960	800	2x 0.465 M
Widescreen 13.5 MHz						
Converted to 1050	26.7 usec	13.5 MHz	11.0 MHz	720	587	2x0.349 M
Progressive systems						
DSC	17.4 usec	36.8 MHz	34.0 MHz	1280	1183	0.922 M
CCDC	17.4 usec	36.8 MHz	34.0 MHz	1280	1183	0.922 M
Widescreen 18.0 MHz						
Converted to 787.5	17.4 usec	27.0 MHz	22.5 MHz	960	800	1.5x 0.465 M
Widescreen 13.5 MHz						
Converted to 787.5	17.4	20.25 MHz	16.5 MHz	720	587	1.5 x0.349 M

* 1/2 of the clockrate used in the compression encoder.
** as specified by the proponent.

MARL

APPENDIX C

PS-WP/4 FINAL REPORT

PS-WP/4 FINAL REPORT

EXECUTIVE SUMMARY

The objective of Planning Subcommittee Working Party 4 (PS-WP/4) was to study and make recommendations regarding the relationship of terrestrial advanced television systems to alternative media, applications and standards. It was also the objective to investigate approaches for growth paths to the future while, at the same time, to support timely decisions on an advanced television (ATV) broadcast system with increased performance quality for the end user. Participants of Working Party 4 have addressed issues related to interoperability, scalability and extensibility and more generally, openness. Representatives of the broadcast television, cable television, program production, motion picture, computer, telecommunications, and imaging industries were active in this working party.

In the prior year's effort (1991), PS-WP/4 developed definitions of key terms such as interoperability, scalability and extensibility. Based upon a world becoming more complex and richer in alternatives (media, transmission/distribution, presentations), the working party developed the concept of image data, defined as the digital equivalent of the video information including image, sound and auxiliary data components. As a result, PS-WP/4 recommended the following in its December 1991 Interim Report:

- Maximize utilization of digital video techniques and image data representation.
- Apply HEADERS and DESCRIPTORS (as agreed by industry standards groups) as a method of identifying image data.

Once the Systems Subcommittee Working Party 4 (SS-WP/4) established the ten selection criteria, PS-WP/4 adjusted its focus to concentrate on the three criteria that related to alternative media: Interoperability, Scope of Services and Features, and Extensibility.

An assessment of the five proponent systems in reference to the above three criteria was made by PS-WP/4. PS-WP/4 developed an OSI-like layered architectural model for ATV to aid in evaluating the proponent systems along with applications and performance questions on these criteria. PS-WP/4 employed a technical consultant, StellaCom, Inc., to assist in this analysis. The assessments were based upon information supplied by each of the proponents in (1) published form, (2) response to specific PS-WP/4 questions and (3) a three-day Interoperability review involving the proponents and a Special Interoperability Review Board (convened specifically for evaluation of the proponent systems relative to the three criteria and conducted in September 1992). The Review Board consisted of experts across a broad array of relevant disciplines. The selected experts had no relationship to any of the system proponents. Results of the Review Board evaluation weighed heavily in the PS-WP/4 conclusions and recommendations.

PS-WP/4 has identified a number of characteristics that contribute significantly to Interoperability, Scope of Services and Features, and Extensibility. These are based on needs and desires exhibited by alternative media advocates, not only for the delivery of terrestrial broadcast television programming but also for other delivery approaches and applications relating to computing, communications, motion pictures and imaging. In relative order of importance, these characteristics are:

- An all-digital implementation based on a layered architecture model
- The use of universal headers and descriptors (as agreed by industry standards group, for example, SMPTE)
- Transmission of the signal in progressive scan format
- Use of a flexible, packet data transport structure
- Viewer transparent channel re-allocation (limited picture and sound while most of the channel capacity is devoted to data transmission for conditional access addressing or other purposes)
- Ability to implement lower-performance, low-cost ATV receivers (comparable price/performance options to current NTSC receivers)
- Ability to implement low-cost ATV consumer VCR
- System architecture and implementation that will allow improvements and extensions to be incorporated as technology advances while maintaining backward compatibility
- Square pixels or at least the option to select square pixel presentation
- Compatibility with relevant international standards or commitment to this objective
- Easily-implementable and user-accessible "still/motion multi-window transmission"

The PS-WP/4 assessment and evaluation of the proponent systems shows some significant differences under the three criteria. Further, all proponent systems need improvement on one or more of the listed characteristics to achieve a desirable degree of interoperability, extensibility and scope of services and features. It is recommended that the Special Panel, the Advisory Committee and the FCC take these differences into account in the process of selecting an ATV standard. Furthermore, selection of a system that incorporates interoperability features not included in the system as submitted for testing requires verification and/or testing. The system submitted for field testing should also include such features.

BACKGROUND

The objective of Planning Subcommittee Working Party 4 (PS-WP/4) was to study and make recommendations regarding the relationship of terrestrial advanced television systems to alternative media, applications and standards, It was also the objective to investigate approaches for growth paths to the future while, at the same time, supporting timely decisions on an advanced television (ATV) broadcast system with increased performance quality for the end user. Participants of Working Party 4 have addressed issues related to interoperability, scalability and extensibility and more generally, openness. Representatives of the broadcast television, cable television, program production, motion picture, computer, telecommunications, and imaging industries were active in this working party.

In the prior year's effort (1991), PS-WP/4 developed definitions of key terms such as interoperability, scalability and extensibility. Based upon a world becoming more complex and richer in alternatives (media, transmission/distribution, presentations), the working party developed the concept of image data, defined as the digital equivalent of the video information including image, sound and auxiliary data components. As a result, PS-WP/4 recommended the following in its December 1991 Interim Report:

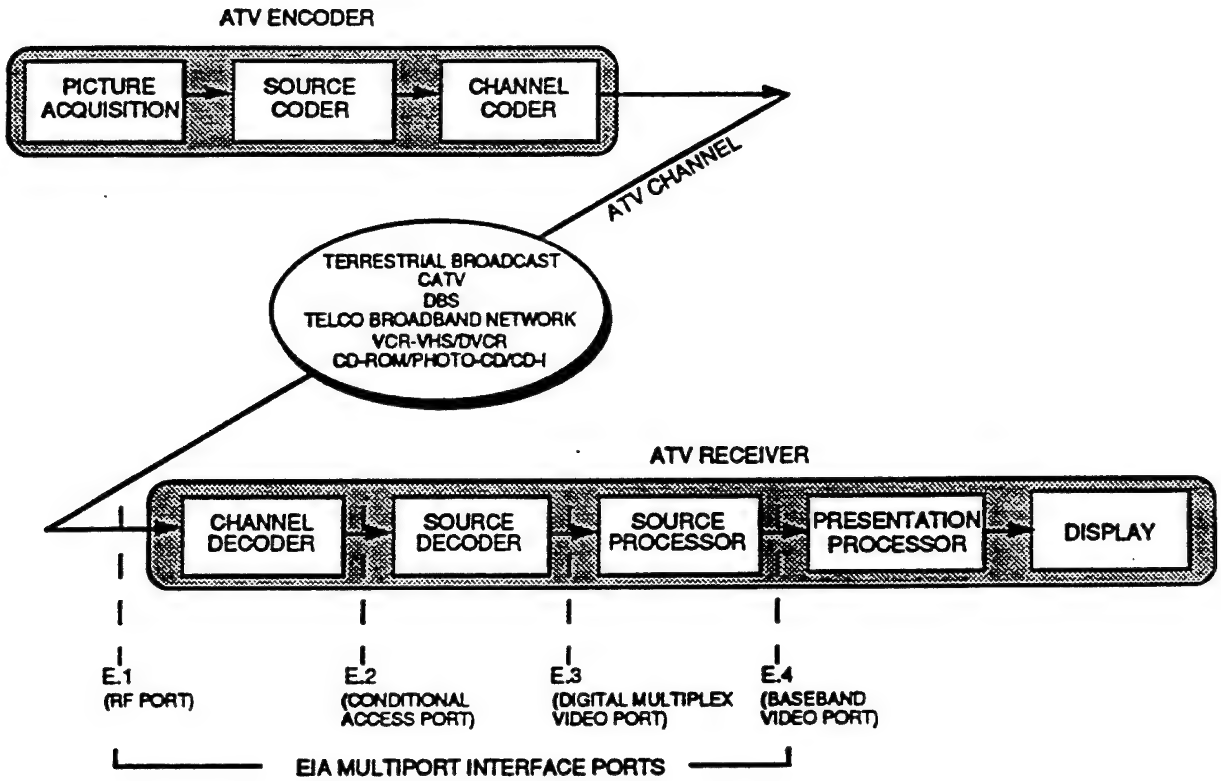
- Maximize utilization of digital video techniques and image data representation.
- Apply HEADERS and DESCRIPTORS (as agreed by industry standards groups) as a method of identifying image data.

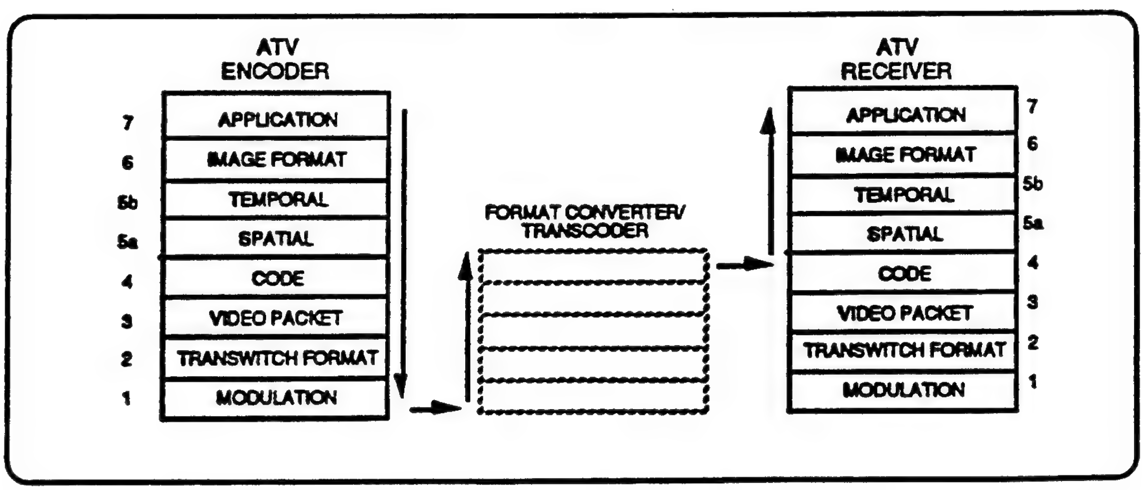
In 1992, PS-WP/4 developed an ISO-like layered architectural model for ATV, consisting of seven layers as shown below:

7	Application
6	Picture Format
5b	Temporal
5a	Spatial
4	Code
3	Video Packet
2	Transwitch Format
1	Modulation

The application of this architectural model in the working party's analyses is shown in the Digital Video Reference Model For Interoperability in the following figure.

ATV REFERENCE MODEL FOR INTEROPERABILITY





Richard Lau
Bellcore
7/9/92
Revised: 11/24/92

PS-WP/4 examined relationships of terrestrial advanced television systems to alternative media, applications and standards in the broad context of recognized and anticipated advances in computing, communications and imaging technology. Weighting the value (e.g., market and social benefit) of interoperability or extensibility other than in established markets (e.g. terrestrial broadcast or cable television programming) was difficult and arguable. Formal positions from affected industries were not available.

Throughout PS-WP/4's investigation, few disagreed that the U.S. FCC ATV decision should endure over some significant, foreseeable technology horizon which should be at least several decades.

Furthermore, few disagreed that over this technology horizon, concepts that are today limited by cost considerations for the consumer market will become reality in the future. These concepts include, at the least, progressive scan (for capture and presentation) and motion image data delivery over packetized wideband networks.

PS-WP/4 believes that the FCC ATV decision must be considered as defining a starting point and assuring a migration path to the future, rather than as defining an end point.

Assessment of Proposed ATV Systems

PS-WP/4 was assigned the task of assessing the interoperability of the proponent systems with media other than terrestrial broadcasting. In addition, the systems were to be evaluated with regard to extensibility and scope of services and features. A preliminary assessment was written in early August by StellaCom, Inc. based on published information available at the time. This report included interoperability questions that were to be answered by the proponents in a planned Interoperability Review in September 1992.

Some 25 generic questions on interoperability, extensibility, and scope of services and features were compiled for all proponents. An additional seven to ten system-specific questions were also addressed to each proponent. Interoperability assessments were made with regard to cable TV, digital technology, headers/descriptors, NTSC, film, computers, satellites, packet networks, interactive systems, format conversion and scalability. The initial use for ancillary data, audio, data, text, captioning, encryption, addressing, capability for a very low-cost receiver with lower performance and VCR capability were assessed as services and features. Extensibility to no visible artifacts regardless of the detail and motion in a scene, to studio quality (editable), to higher resolution such as VHDTV and UHDTV and provision for future compression enhancements were examined.

The proponents answered all questions in writing at the Interoperability Review on September 23-25. They were given additional time to submit any proposed interoperability improvements to their systems. A revised interoperability assessment was then generated including the updated material. The proponents were given the opportunity to critique the

assessment. The assessments were then revised and abbreviated to five to six pages for each proponent to be submitted to the ACATS SS-WP/4 Special Panel. The result is attached as Reference 1.

Interoperability Review

An interoperability Review was held with the ATV proponents to focus on interoperability, extensibility and scope of services and features. Two-hour presentations were made by each proponent followed by 1/2 hour of questions from the Review Board of multi-industry experts and 1/2 hour by other PS-WP/4 members.

The Review Board members were:

Jules Bellisio	Bellcore
Jules Cohen	Consultant/MSTV
Gary Demos	Demographx
Jack Fuhrer	Hitachi America
Branko Gerovac	DEC
Jukka Hamalainen	Matsushita Applied Research Lab
George Hanover	Electronic Industries Association
Robert Hopkins	ATSC
Ed Horowitz	Viacom
Mike Liebhold	Apple Computer
Craig Tanner	CableLabs
Arpad Toth	Eastman Kodak
Tony Uyttendaele	CapCities/ABC

The proponent presentations in the Interoperability Review gave strong consideration of interoperability, extensibility and scope of services and features. The proponent ATV system descriptions contained evidence of effort toward enabling interoperability with alternative media. The written material distributed at the meeting is found in Reference 3: Minutes of the PS-WP/4 Interoperability Review, September 23-25, 1992.

Additional information was submitted by some proponents on October 1, 1992 and can be found in Reference 4: Minutes of the PS-WP/4 meeting, November 5, 1992. Evaluations from the Interoperability Review were submitted to StellaCom, Inc. by members of the Review Board for compilation and analysis. As much as possible, the evaluations were based on "what the evaluator expects the proponents will deliver commercially". The proponents were also invited to evaluate the proposed systems. The results from these evaluations are attached as Reference 2. It contains summaries of the ratings by the proponents, the Review Board and the ratings by individual experts.

Conclusions

Based on assessments of the five proposed ATV systems and the Interoperability Review conducted on September 23-25, 1992, PS-WP/4 reached the following conclusions with respect to the selection criteria of Interoperability, Scope of Services and Features, and Extensibility. See Reference 2 for numerical ratings and detail.

Overall

During the assessment and review process, all proponents gave strong consideration to meeting the three selection criteria and presented evidence to show how their systems met or could meet the criteria.

The four all-digital systems (AD-HDTV, CCDC, DigiCipher and DSC-HDTV) were found to be superior to the Narrow-MUSE system with respect to Interoperability and Extensibility on all but two or three characteristics where Narrow-MUSE was approximately equal to the all-digital systems. With respect to Scope of Services and Features, the all-digital systems also ranked better than Narrow-MUSE but the difference was smaller.

Interoperability

- All five systems were judged to be easily deliverable by satellite or Cable TV.
- The four all-digital systems were judged to interoperate well with NTSC.
- The four all-digital systems were judged to interoperate well with film with CCDC ranking better than the other three systems.
- Interoperability with computers was judged to be more difficult than with satellites, Cable TV, NTSC and film. DSC-HDTV ranked slightly better than CCDC, with AD-HDTV next, and DigiCipher ranked somewhat worse.
- Progressive scan and square pixels are important for computer and other image applications (rotation, scaling, still-image, LCD displays, ...). Computer industry applications desire greater than 70 Hz display rates although no proponent system offers this feature.
- All of the proponents of all-digital systems now recognize that the use of headers and descriptors is a critical enabling concept for assuring ATV flexibility in the future. Although these four proponents propose to include headers and descriptors in their systems, only AD-HDTV had its final proposal for headers and descriptors fully implemented at the time the system was tested by ATTC, and this system received the highest rating on this characteristic. There is need for definitive work on an industry standard for headers/descriptors. The SMPTE work is a candidate.

- All of the proponents of all-digital systems now recognize that the use of a packetized data structure combined with headers and descriptors is important for future network and telecommunications applications of ATV. Although these four proponents propose to use a packetized data structure, only AD-HDTV had its current proposal for a packetized data structure implemented at the time the system was tested by ATTC and this system received a significantly better rating than the others on this characteristic. The DSC-HDTV system, as tested, incorporated a form of packet structure using fixed-length data segments.

- Interoperability with interactive systems is an important consideration for future applications. The four all-digital systems were rated approximately equal with respect to this characteristic.

- With respect to format conversion, the four all-digital systems were rated approximately equal.

- Scalability is an important characteristic in that it will permit user price/performance tradeoffs at both the display (picture) level and at the network (transport) level. All of the systems are scalable at the picture level, permitting display quality and cost to be traded over a wide range. None of the systems achieve the degree of scalability at the transmission level that would permit similar trade-offs in future 'bandwidth-on-demand" network environments and low-cost receivers. However, the, AD-HDTV and DSC-HDTV systems both employ a prioritized data structure at the transmission level that provides a limited degree of scalability.

Scope of Services and Features

- With respect to the nine characteristics considered for this selection criterion (Initial use for ancillary data, Audio, Data, Text, Captioning, Encryption, Addressing, Low-cost receiver and VCR capability), there were no statistically significant differences in the ratings of the four all-digital systems.

- The use of a packetized data structure with headers and descriptors as discussed above provides important system flexibility in meeting this selection criterion. It allows channel capacity to be dynamically assigned to services and features as needed, and it permits new services and features to be implemented in the future, while maintaining backward compatibility with existing ATV receivers. Computer, telecommunications and cable TV industry experts place a high value on "viewer transparent channel data reallocation", permitting instant reallocation of the majority of the channel capacity to data transmission while maintaining audio transmission and limited video transmission, e.g., a still image or motion video (possibly at reduced quality) in a window or windows on a still background.

- With respect to low-cost receivers, this characteristic was modified during the review to mean the ability to implement an ATV receiver with reduced performance that would sell at a price comparable to current NTSC receivers. SS-WP/3 is addressing this issue, and current indications are that there is no significant difference among the five proposed systems, and they were rated essentially equal with respect to this characteristic. Further, the current view is that signal-processing electronics cost will ultimately become an insignificant portion of the total receivers' cost.

- The use of a packetized data structure with headers and descriptors combined with a prioritized data structure at the transmission level is an advantage for accomplishing trick modes in digital VCRs. Special flags could allow easy and low-cost implementation for trick play and editing.

- The use of a packetized data structure with headers and descriptors allows flexibility in implementing audio services. Independently coded audio channels can be used when appropriate (e.g., for simultaneous multiple languages during a newscast or a sportscast), and compositely coded audio can be used to provide five-channel "surround-sound" for major movies. The required data capacity will vary depending on the type of coding and the number of channels.

Extensibility

- With respect to the six characteristics considered for this selection criterion (Extensibility to: no visible artifacts; studio-quality data rate; higher resolution; VHDTV; and UHDTV; and Provision for future compression enhancements), there were no statistically significant differences in the ratings of the four all-digital systems.

- The use of a packetized data structure with universal headers and descriptors as discussed above, provides important flexibility in meeting this selection criterion. For example, if a higher data rate channel is used to distribute programming to network affiliates, additional packets (with appropriate headers and descriptors) could provide higher-quality images for post-production processing by the affiliate. For simple pass-through, these packets would be stripped, and the remaining data stream re-clocked to the terrestrial broadcast data rate.

- The proponents of the four all-digital systems have all made provision to incorporate future improvements in their video compression systems that will be backward compatible with existing ATV receivers. Examples of possible improvements include more accurate motion estimation, better psycho-physical decision rules, improved buffer management rules, etc.

System Specific Comments

During the assessment and review conducted by PS-WP/4, the proponents described how they proposed to meet the requirements of the three selection criteria against which they were being evaluated. In some cases, these proposals require changes to the system hardware and software from that which was tested at ATTC. The effects these changes could have on the results of the ATTC and ATEL evaluations of the systems are unknown. In other cases, certain features or characteristics were present in the system when it was tested by ATTC, but were not exercised during the tests. Where changes are required, it must be verified that these changes will not affect the results obtained by ATTC and ATEL. Where the features were not exercised, it must be verified that they work when exercised.

All proponent systems need some improvement in interoperability, scalability and extensibility. For example:

AD-HDTV

- verification of the progressive-scan and square-pixel migration strategy

CCDC

- implementation of universal header/descriptor concept
- implementation of packetized data structure

DigiCipher

- implementation of universal header/descriptor concept
- development & verification of a progressive-scan and square-pixel migration strategy
- implementation of packetized data structure

DSC-HDTV

- implementation of universal header/descriptor concept
- complete the implementation of a fully packetized data structure

Narrow-MUSE

- commitment to digital implementation
- implementation of universal header/descriptor concept
- implementation of packetized data structure
- development & verification of a progressive-scan and square-pixel migration strategy

RECOMMENDATIONS

Based upon the needs and desires exhibited by the alternative media, not only for the delivery of terrestrial broadcast television programming but also for other delivery approaches and applications relating to computing, communications, motion pictures and imaging, PS-WP/4 has identified a number of characteristics that contribute significantly to Interoperability, Scope of Services and Features, and Extensibility. In relative order of Importance, these characteristics are:

- An all-digital implementation based on a layered architecture model
- The use of universal headers and descriptors (reference ACATS Interim Report, March 1992)
- Transmission of the signal in progressive scan format
- Use of a flexible, packet data transport structure
- Viewer-transparent channel re-allocation (limited picture and sound while most of channel capacity devoted to data transmission for conditional access addressing or other purposes)
- Ability to implement lower-performance, low-cost ATV receivers (comparable price/performance options to current NTSC receivers)
- Ability to implement low-cost ATV consumer VCR
- System architecture and implementation that will allow improvements and extensions to be incorporated as technology advances while maintaining backward compatibility
- Square pixels or at least the option to select square pixel presentation
- Compatibility with relevant international standards or commitment to this objective
- Easily-implementable and user-accessible "still/motion multi-window transmission"

The PS-WP/4 assessment and evaluation of the proponent systems shows some significant differences under the three criteria. Further, all proponent systems need improvement on one or more of the listed characteristics to achieve a desirable degree of interoperability, extensibility and scope of services and features. It is recommended that the Special Panel, the Advisory Committee and the FCC take these differences into account in the process of selecting an ATV standard. Furthermore, selection of a system that incorporates interoperability features not included in the system as submitted for testing requires verification and/or testing. The system submitted for field testing should also include such features.

APPENDIX D

SMPTE HEADER/DESCRIPTOR TASK FORCE: FINAL REPORT

SMPTE Header/Descriptor Task Force: Final Report

January 3, 1992

This report of the SMPTE Task Force on Headers/Descriptors is an approved document of the SMPTE Standards Committee and is made available for information, as it contains valuable proposals concerning the development of digital imaging and video systems and for standardization of certain of their aspects that will be of interest generally. The standardization aspects of the report will be further considered under the normal processes of the SMPTE for the creation and approval of engineering documents, which includes the opportunity for further comment and for public review prior to their final acceptance. Persons wishing to actively participate in the development of these standards, including attendance at Working Group meetings and ballot response, may contact the Engineering Dept. of the SMPTE. It should be noted that engineering documents arising from the contents of this report may differ significantly from its recommendations, and caution is suggested in the use of this report as the basis of design or implementation.

1.0 Introduction

The Task Force on Header/Descriptors has considered the questions posed in its scope of committee work and makes the following final report and recommendations to the Standards Committee. The report begins with a discussion of the general objectives of the header/descriptor, and then presents more specific objectives selected by the Task Force as it developed two alternative implementations.

Both of the proposed implementations could support new SMPTE standards, and are described in some detail here. The "ASN.1 Implementation" is structured using only Abstract Syntax Notation 1 (ASN.1), an existing and evolving ISO/CCITT standard principally used in the computer industry. The "Compact Implementation" is designed to minimize the number of bits allocated to the header/descriptor function, but also permits optional use of the ASN.1 notation later in the header/descriptor for further extensibility. Both implementations perform essentially identical functions.

Appendices A and B present illustrative approaches to the design of transport headers and header-decoding software, respectively. Transport headers are designed to address certain difficult data transport problems. Appendix C lists the official task force members as of January 3, 1992.

In view of (1) the great importance to industry and its customers of the capabilities provided by the header/descriptors described below, and (2) the degree to which these two possible implementations satisfy the objectives established at the outset for header/descriptors, the Task Force recommends that: the Standards Committee arrange for the preparation of one or two new standards for digital header/descriptors based on either the "Compact" or the "ASN.1" Implementations described below, or on a combination thereof.

2.0 General Objectives

The header/descriptor task force was directed to consider header/descriptor architectures and implementations appropriate for the emerging digital high-definition television (HDTV) and high-resolution system (HRS) industries. The primary design objectives of the task force are:

- *Universality.* All image and other data streams should be labeled so that signals can be shared across systems and applications with minimal degradation or confusion: the header/descriptor should therefore uniquely identify the encoding scheme employed and how the data is to be interpreted.
- *Longevity.* The header/descriptor should provide a number of potential identification codes adequate to serve for decades, and preferably centuries: this implies that specific encoding identifiers, once assigned and registered, should not be reassigned or redefined. The header/descriptor should also facilitate longevity for equipment and media of all types.
- *Extensibility.* To facilitate service enhancement and innovation, and to promote longevity of both equipment an recorded signals, the header/descripto should accommodate technological ad vances in either equipment or recorde signals with minimal risk of obsoleting ex isting components, infrastructure, and me dia collections.
- *Interoperability.* The header shoul permit optimal sharing of data stream across data-generation, carrier, and equip ment technologies and services in a variet of error environments and should permi all equipment and applications to success fully ignore encrypted or otherwise delib erately inaccessible data.
- *Cost/performance effectiveness.* Th header/descriptor should permit use o both low-cost equipment as well as mor expensive high-performance equipment the header/descriptor should also accom modate inexpensive equipment incapabl of decoding all possible data stream Economy and simplicity through flexibilit and scalability of the key performance pa rameters should also be supportable.
- *Compactness.* The header/descripto should be economical in its utilization o bits and should typically comprise a negli gible fraction of the underlying dat stream.
- *Rapid capture.* Much video and othe serial data is intercepted midstream, suc as when users switch to a new channel, an therefore the header/descriptor shoul permit rapid header identification, ade quate to meet the needs of all applica tions.
- *Editability.* Common editing and pars ing operations, such as splicing, append ing, replacing, inserting, cropping, and overlays, should be supportable by the header/descriptor architecture without necessarily requiring decoding and encod ing of the data stream itself.

3.0 Specific Header/Descriptor Objectives

To meet the general objectives summarized above, the Task Force selected the following compact set of specific objectives which are met by both implementations described later. The header and descriptor are defined here separately.

3.1 Specific Header Objectives

The specific objectives of the header are to:

• Identify by number the encoding standard employed by the attached block of data.

• Specify the length of that block of data, so that equipment of any epoch can successfully skip uninteresting blocks of data or data encoded using standards defined subsequently.

• Indicate whether a readable descriptor follows the header.

• Permit users to intercept data streams at random times, as when switching channels, so that proper data interpretation begins swiftly.

• Provide optional error-protection capability. Data generation entities may wish to supplement error-protection services provided in subsequent environments experienced by that data, particularly when those environments are unknown.

The task force considers these attributes of the header to be the minimum mandatory set, recognizing that additional important capabilities can be provided by the descriptor.

3.2 Examples of Header Use

A simple example illustrates how these minimal capabilities for the header satisfy the general objectives discussed above. Suppose, after many years, some HDTV broadcasters wish to provide dual-language sound tracks. This capability could be provided by adding to the data stream blocks of data conveying the second language. These new blocks would be labeled by a header incorporating a standard number not recognized by equipment produced earlier. This older equipment would read the header and recognize the standard identification number as being unknown. It could then observe the length of the associated block, and skip over it to the next header.

All data could be labeled by such flexible headers, or only a designated portion (e.g., "auxiliary data") of a more rigidly defined larger video data stream. Note that HDTV receivers capable of receiving only 20 Mbits/sec could not accommodate increases except at the expense of any spare capacity previously reserved for expansion, or by the broadcaster reducing the number of bits conveying video or audio; in the latter case the original standard would have to be defined so as to permit receivers to accommodate any such real-time video or audio truncation, however.

3.3 Specific Descriptor Objectives

The principal function of the descriptor is to convey additional information that improves the usefulness of the data to the user; its format would be specified independently of the standard employed for the data itself. Such optional auxiliary information in the descriptor might include transport information such as cryptographic, priority, or additional error-protection information, as well as source time, authorship, ownership, restrictions on use, royalty payment information, explicit description of encoding or decoding processes, intermediate processing performed, and other information in forms that could evolve over the years. To simplify the decoding task, the descriptor may also contain an abbreviated table of contents and a flag indicating whether any information has changed since the previous descriptor. The beginning of the descriptor would also indicate the descriptor length so that it might be skipped without interpretation if the user chooses. Optional additional error protection would be available for data originators so desiring it.

Specifically, the descriptor could include:

• A list of standard-identification numbers, parameters of operation, text, and algorithms, in any desired combination.

• A compact optional table of contents for the descriptor.

• A flag indicating whether changes occurred since the previous descriptor.

• The length of the descriptor so that it might be readily skipped if desired.

• Information indicating the number of descriptor entries and their formats so that they might be properly interpreted.

• Optional error protection for the descriptor.

The presence or absence of a descriptor could be indicated by one of the bits contained in the header.

3.4 Examples of Descriptor Use

The use of standard identification numbers in the descriptor permits very compact and flexible encoding. For example, one such number might be allocated internationally to each model number of studio television camera, so that subsequent image processing can maximally improve image quality, compensating for any camera idiosyncrasies. Similar identifiers could be used for different forms of physical, analog, or digital filtering that has been applied to the image subsequently, so that user equipment might again appropriately refilter the image in an optimum way depending on the user's intentions. This is important because performance for any particular display device or audio system is best when it reflects the processing that has occurred previously.

Authorship, ownership, and other such information could be conveyed by compact standard-identification numbers, or by use of plain text in English or another language. Certain descriptors might simply be numbers indicating the settings of certain switches at the time the signals were generated, such as switches controlling audio bass, treble, or volume. The descriptor may also include subroutines or other encoded instructions that facilitate subsequent processing or decoding.

Standards numbers and parameter fields can also be used to support transport-layer functions, including essentially all forms of cryptography, statement of the relative priority of the current data block, priority "bidding" data (so users can bid for priority in a free-market sense), synchronization reinforcement blocks, and other information, the character of which can be defined over the years as new standards identification numbers are assigned and as new languages and protocols are defined. The incorporation of transport capabilities in this standard should not compromise those established by other layers, but would merely supplement them. The only limitation is that such descriptor transport standards should remain robust if block sequences are shuffled in another transport layer.

4.0 Illustrative Examples of Header/Descriptor Use

An HDTV broadcaster could simply divide the HDTV signal into blocks, each beginning with a header of perhaps 6 to 16 bytes in length. This header would contain the length of each block, which could be fixed for all time or variable, and a unique standards number indicating the given HDTV encoding protocol, which may also be unchanging in the initial years. If the over-the-air broadcast standard is heavily error protected, little additional error protection might be added to the header. Good engineering practice would suggest, however, that the header be independently error protected using some of the options described later, and that separate, and possibly less robust, error protection be applied to the remainder of the data stream.

If the HDTV channel is defined so as to perform all transport functions, including all synchronization and error correction, then the header/descriptor described here might be embedded in the transported data stream. At that level it would preferably be used to encapsulate all data, but could be used in an inferior implementation to encapsulate and characterize only substream or side-channel data. In this example too, the descriptor might convey the origins and processing history of the data, enabling future higher-performance systems to employ post-filters optimizing the quality of the output images. Such flexibility could be particularly important if the output display capabilities of the equipment enabled flexibility in frame rate, pixel interpolation routines, and chrominance manipulation. In the case of audio signals, descriptors could be used in a similar manner to enable optimum reproduction by characterizing microphone placement and preprocessing.

A variety of header/descriptors plus associated data could be sequenced in any side channel to provide a variety of flexible delivered products, such as multiple audio channels, multiple-language captioning,

home-shopping order information, and even entire TV signals encoded at lower quality so that more than one can share a single channel. In this example, the ultimate flexibility that could be obtained with individual broadcast channels would be heavily dependent upon the flexibility inherent in the broadcaster's ability to decrease the data rate associated with the initial broadcast product to accommodate possible growth in use of the side-channel capability.

It is also possible to send descriptor information in a separate block with its own header, rather than embedding it in the same blocks as the associated data. In this case the characterization of a signal provided by a descriptor could remain unchanged for many blocks until that descriptor information changes. In the event the transport layer is prone to shuffling the sequence of blocks, impairing association of descriptors and use of this data, the inherent flexibility of this header/descriptor system permits incorporation of sequence numbers in blocks so that the receiving entity can properly sequence them. Such separation of descriptor information into separate blocks also simplifies translation between environments having different transport layer protocols; any descriptor elements providing transport-layer functionality can be added or deleted when moving between such environments, as desired. An option whereby special "transport header/descriptors" are prepended to blocks is described in Section 5.2.7.

To the extent that transport-layer functionality might be embedded within the basic data block, it could be harmlessly overlaid by similar transport-layer functionality in other system elements without loss. Thus the great flexibility inherent in the header/descriptor architecture proposed here, including its ability to perform multiple functions in a multilayer ISO environment, should not be a handicap, and may in some applications be an important advantage.

Furthermore, this header/descriptor structure permits efficient utilization of the standards activities of a wide variety of national and international organizations. Every standard developed by such bodies for characterizing or communicating digital information can be characterized by an identification code that can be conveyed in an efficient manner by the header/descriptor system described here. Since the implementations proposed here call for each standards authority to have its own identification number, such standards bodies could, in this "ID" concept, choose to use the very same identification numbers they have previously chosen for other purposes. Alternatively, that authority could choose some other simple one-to-one mapping between numbers. At the same time, such standards can also use this header/descriptor system as an embedded construct within their own protocols. Although the full power of the flexible extensible structure proposed here will only become apparent as it is developed and improved over the years, the basic architecture can be fixed immediately. This would permit immediate fabrication and utilization of equipment based upon this standard.

5.0 Compact Header/Descriptor Approach

5.1 Compact Header/Descriptor Architecture

The header is divided into two parts: a 2-byte "header key," and the remainder, or "header tail." Among the available header options are those without tails, corresponding to a 2-byte header conveying simple messages: 32 possible messages of this type are available for future definition. The most usual function of the key, however, is merely to "unlock" the tail by providing information about the format of that tail.

The descriptor is divided into three parts: a 2-byte "descriptor key" (similar to the header key), a "core" of 0 to 8 bytes, and the "descriptor tail," which conveys a series of numbers signifying various pieces of information.

5.2 Header Key

5.2.1 Header Key Organization

The header key consists of two 4-bit fields and one 8-bit field. The 8-bit field provides 2-bit error-correction capability for the key, and the first 4-bit length-type "LT" field determines the length of the header field in the tail that contains the length of the block. A block of data is defined as comprising the header, any descriptors, and the associated data. The other 4-bit "ID" field normally determines the number of bits in the header-tail field devoted to indicating the standard number under which the remainder of the data in the block, possibly including the descriptor, is encoded. Most descriptors would be publicly readable, however. The ID field also indicates whether a readable descriptor follows the header; sometimes one might wish to read the descriptor before deciding whether to decode the rest of the data block. The ID field can alternatively convey messages for certain 4-bit combinations in the first 4-bit LT field. In addition to specifying the length of the header tail field devoted to specifying block length, the LT field also determines what level of error protection is being provided for the header.

5.2.2 Header Key: 4-Bit "Length-Type" LT Field

The primary purpose of the 4-bit LT field is to specify the length of the field in the header tail devoted to specifying the block length. For 6 of the 16 possible combinations proposed here (the "fixed-length" options) the total length of the block is prespecified so that no bits in the header tail are allocated to specifying block length. The other 10 proposed combinations permit use of 1, 2, 4, and 6 bytes for specifying the block length in bytes in integer format; in each case versions with and without error-protection capability for the header are available to the standards definition community. The 8-bit protection provided to the header key is always present, however, since the integrity of the key is crucial, and it represents such a small part of the total.

The six proposed fixed-length options are as follows:

1. The block is 2 bytes long and the message is conveyed by the 4-bit ID field: 16 messages are possible. One of these messages could signify that the rest of the header/descriptor is coded in ASN.1 (dual compact- and ASN.1-header decoding capability would be necessary for all equipment, however, but Appendix B suggests this might not be burdensome).

2. Same as (1), except that an additional 16 messages are possible (for a total of 32 2-byte options).

3. The block is 4 bytes long, the message consisting of the 4 ID bits plus 2 bytes, a total of 20 bits.

4. The block length is 6 bytes: 28 bits $(4+3\times 8)$ are available, the remaining byte providing error protection for the last 4-byte set.

5. The block length is 6 bytes with 36 bits of information $(4+4\times 8)$ being available, but without additional error protection.

6. The length of the block is unknown or irrelevant.

In each of the foregoing options, except the first two, the available bits can be divided in a yet-to-be-determined way between those indicating the standard identification number and any additional message. The proposed SMPTE standard would constrain only options 4 and 5, allocating the first 8 bits to designating the sovereign state, so that development of these 6-byte options can proceed without international agreement. Designation of standards bodies and sovereign states is discussed further in Section 5.2.4.3.

The ten remaining proposed options for the LT field provide for either 1, 2, 4, or 6 bytes in the header tail to be allocated to specifying the block length in bytes, always in integer format. Depending on which of these ten options is chosen, additional information is also conveyed concerning the level of error protection for the header. These proposed options are as follows:

7. One-byte field in the header tail specifies block length in integer format; no additional error-correction capability is provided to the header. Blocks up to 256

bytes long are available under this option.

8. Same as (7), except an additional 1 byte is provided for header error protection.

9. Same as (7), except that 2 bytes specify block length: the maximum block length here is 64 Kbytes.

10. Same as (9), with an additional byte for header error protection.

11. Same as (7), except that 4 bytes specify block length, which can approach 4 billion bytes.

12. Same as (11), except that 2 bytes of header error protection are added.

13. Same as (7), except that 6 bytes specify block length.

14. Same as (13), except that 2 bytes of header error protection are added.

15–16. To be determined.

5.2.3 Header Key: 4-Bit ID Field

When the 4-bit ID field is not being used to convey a message using LT options 1 to 6, then the first 3 bits of the ID field indicate one of eight possible lengths for the header-tail field conveying the standard identification number, and the fourth ID bit indicates whether a readable descriptor follows the header. These proposed eight tail-length field options include 1, 2, 4, and 8-byte options. These eight tail-length options proposed for the ID field are as follows:

1. One-byte standard identification number allocated internationally.

2. Same as (1), but providing for 256 additional possible standards.

3. Similar to (1), except that the standard field in the header tail contains 2 bytes instead of 1, providing for 64,000 international standards.

4. Similar to (1), but with a 3-byte field in the header tail, providing for over 16 million international standards.

5. Two bytes in the header tail indicate the standard identification number, the first byte indicating the sovereign state (see Section 5.2.4.3).

6. Similar to (5), except that 4 bytes are available (one for the sovereign state).

7. Similar to (5), except that 8 bytes are available.

8. To be determined, or reserved for the distant future.

These first four options provide 1, 2, and 3-byte standard identification numbers to some designated international standards body or bodies. Two versions of the 1-byte option are provided because otherwise too few combinations would be available. The remaining options provide 1 byte (or more) that identifies the sovereign state under whose authority the remaining standard-identification-number bytes the ID field have been assigned (1, 3, or 7 bytes remain). Note that receiving equipment would generally not distinguish between sovereign state identifiers and standard identification numbers: they would be treated together only as a single merged number that was known or unknown.

5.2.4 Header Tail

5.2.4.1 Header Tail Organization

The header tail contains one field for indicating the block length in integer format, one field for the standard identification number, and optional fields for error protection, all having been discussed above.

5.2.4.2 Header Tail: Standard ID Organization

The Standard ID comprises two parts: the sovereign identification number, and the standard identification number, described in Sections 5.2.4.3 and 5.2.4.4, respectively.

5.2.4.3 Header Tail: Sovereign State Identification

Eight bits is sufficient to designate the authorizing sovereign state, even if the number of sovereign states exceeds 256: the trick is to subdivide certain undesignated sovereign state identifiers by borrowing bits from the remaining standard identifier field. For example, by deferring to the United Nations the task of defining sovereignty, it should be an easy matter to assign all present U.N. members unique identification numbers in alphabetical order, and to assign the next numbers in order of membership admission to states not replacing member states who are already members. Once 224 member states exist, new U.N. members would share the last 32 numbers. Within these 32 sovereign state "condominium" identifiers, an additional 3 bits would be borrowed from the Standard Identifier field, permitting 8 new member states per condominium, or $256-32+(32\times 8)=480$ possible states. Of these state designators, 32 would be assigned to existing standards bodies. In the unlikely event more states are created, still more bits can be borrowed, permitting unlimited growth.

5.2.4.4 Header Tail: Standard Identifier

The Standard Identifier would be selected by the indicated Sovereign State or International Standards Body using existing procedures. To the extent standards already have unique identification numbers, those same numbers could be used here. To the extent they do not, each standards body should map their standards into numbers, preferably a compact or consecutive set. Most new standards might be introduced under long numbers associated with minor standards bodies, while standards achieving wide acceptance could be renamed with short n... ...e bits can be borrowed from the Standard Identifier field almost indefinitely, the system would permit in certain cases (the longer header options) for every individual ever born to become a sovereign definer of standards under some state's authority (the 8-byte option provides each sovereign state with 70 million billion numbers). In the same way, a standards body could allocate numbers well spaced numerically so that their least significant bits could become "user-allocated bits," functioning in lieu of, or in addition to, a descriptor.

Note that the cost of this kind of flexibility, as well as the other sorts of flexibility described earlier, is essentially negligible given the extensible nature of this header architecture. That is, the price of using long headers is paid principally by those who need them, while most users will prefer and use the shorter options.

5.2.5 Header Organization Illustrations

These options can be represented pictorially. The 16-bit composition of the header key consists of three parts: the 4 bits in the length-type LT field (l), the 4 bits in the ID field (i), and the 8 core error-protection bits (p): these protect only the key. The header key always consists of: *llllliiiipppppppp*.

If we designate this 2-byte key by the symbol *K*, and the bytes representing the block length field by *L*, the bytes representing the standard identification number by *S*, and the header parity bytes by *P*, then for LT options 1 and 2 we have only *K* (2 bytes). For LT options 3, 4, and 5, we have only *KSS* (4 bytes), *KSSSP* (6 bytes), and *KSSSS* (6 bytes), respectively. In this case of extremely short blocks, the standard identification numbers *SSS* and *SSSS*, combined with the 4 bits in the ID field, would generally convey messages in their own right, and could also be used for a variety of transactional and transport functions. LT option 7 would permit the following possibilities: *KLS*, *KLSS*, *KLSSS*, *KLSSSS*, and *KLSSSSSSSS*, where the number of bytes allocated to the Standard Identification number by *S...S* is specified by the 3 bits in the ID field (see 5.2.3). One likely option for HDTV might be LT option 12 combined with ID option 1, represented as *KLLLLSPP* and comprising 9 bytes. The standard number contained in field *S* would be an international standard. If such an international standard were devised to have block lengths no larger than 64 Kbytes, and if 1-bit error correction were adequate, then a 6-byte HDTV option is available: *KLLSP*. Such short 5 or 6-byte header options could often be employed for nonvideo information. For example, *KLSS* (5 bytes) would accommodate 64,000 possible international standards employing data up to 256 bytes per block.

5.2.6 Header Architecture: Equipment Implications

Equipment interpreting such headers can be particularly simple. Since it normal-

ly would be reading a stream of blocks, and should know when a block begins, it could simply jump to one of the 64K words specified by the 16-bit header key, these words indicating the location of the bytes specifying the block length. An example of an algorithm to perform these header decoding functions appears in Section 5.2.8, and an illustrative program coded in C appears in Appendix B. Simpler equipment might simply look at the first 4 bits of the LT field, from which the same information can be deduced in an error-free environment. Equipment of intermediate complexity can use the key error-protection bits to an intermediate degree. The ID bits immediately indicate the field where the Standard Identifier is located, and equipment should compare this to a list of standards it is prepared to process. After any indicated processing, the equipment moves directly to the next header at the specified number of bytes along the data stream.

Should synchronization be lost, it could readily be recovered in most cases of interest. For example, in the 9-byte HDTV header example above, these 9 bytes would probably never change within a single broadcast program. HDTV broadcast receivers would simply scan the data stream for that particular 9-byte sequence, which could be used as a traditional synchronization block. This would work even if multiple header types were being interleaved. Accidental synchronizations (very rare) in the data block would be recognizable because the indicated false block length would probably not lead to a valid header.

The most direct use of these header/descriptors would be as a series of bytes at the start of each data block, where the data blocks are then concatenated in a comma-less string of bits. Other uses could also be made, however. For example, a particular transport scheme (e.g., over-the-air HDTV) might package data blocks in discontinuous but fixed form within a larger data stream, this stream being characterized by the embedded header/descriptors. A header/descriptor could also characterize (say) each frame of an HDTV sequence, and the same header/descriptor technique could also be used within each frame to communicate details about its internal structure. Such nesting of header/descriptors does not impair synchronization much, although an initial false synchronization could occur within a larger block; this would be quickly detected if the presumed block ended without a valid header following it. The search for a valid header could then resume.

5.2.7 Header Architecture: Transport Functionality

If one wishes to incorporate synchronization augmentation, extra error protection, block prioritization information, or cryptography for the header or descriptor, it must either be defined at the outset in the header/descriptor definitions, or its incorporation becomes standard-specific — but how could we know the standard if we were not synchronized or error-free, etc.?

For example, in high-error-rate environments users may wish to provide their header/descriptors with more synchronization information or error protection than is available in the basic definitions. Fortunately, much additional transport-related information can be conveyed efficiently through repetition of short blocks. One simple approach is to insert brief bursts of short 2-byte headers into the data stream, where the synchronization powers of these 2-byte header blocks are cumulative: the greater the anticipated channel noise, the longer these bursts should be. Although the software of some receivers may not be sophisticated enough to synchronize such bursts well, this approach is standard independent, and so such software could be designed today. Such a burst inserted anywhere in a sequence of blocks permits synchronization of the entire stream. Since 2-byte headers can correct 2 bits in 16, synchronization bursts fail only when the bit-error rate continually exceeds ~0.2. In still higher noise environments where it is known such 2-byte bursts might be employed, the receiver can either autocorrelate the signal or correlate it with potential 16-bit synchronization words; this could provide synchronization for nearly any bit-error rate, provided the burst length was sufficiently long.

Similar issues arise when extra error protection for the header/descriptor is desired. If the nature of such protection is defined only in a data or standard ID area that cannot be read unless error protection is employed, the data is generally inaccessible. One option is for the user to generate, as above, bursts of short blocks that convey primarily the identity of the desired error-protection scheme employed in the longer data blocks. Such standards could be defined so that they are presumed operative for all following data until "turned off" by another command. Alternatively one of the short blocks (say 2 bytes) could be employed to indicate that error protection, cryptography, or other such schemes were to be conveyed in a "turn-on, turn-off" fashion; a separate short block could be used to convey the opposite message. Such a repetitious series of short blocks can also be well protected and synchronized in essentially any reasonable error environment (BER <0.2) using the existing defined header/ descriptor options. In this case the user would need to know only the definition of the chosen error-protection ID number. Such error-protection schemes and protocols can be defined in simple ways over the years.

Yet another similar problem involves the possible incorporation of block priority information. For example, if some data processing, transmission, storage, or display step cannot handle all the data, the originator of that data might wish to tell the user which data is more expendable. Yet the user might wish to do only minimal decoding to determine priority. Although such information would then be descriptor-standard-specific, one or more descriptor standard IDs could be defined in such a way as to convey relative block priority, say on a scale from 1 to 10. Such information could also be embedded in the data itself, but this would generally require the user to do more decoding before discarding any block. Although confusion could arise because multiple priority-labeling descriptor schemes might arise, they would all have standard numbers which, in principle, could be accessed. Section 5.4 describes another approach, which can be used in parallel.

Although this discussion has not been exhaustive, it does suggest that the proposed header/descriptor definition has great flexibility for handling a variety of problems faced when it must supplement or provide transport-layer or other OSI layer functions.

5.2.8 Compact-Header Decoding Algorithm

For simplicity here, we assume the block has been synchronized and that the header key has been error-corrected, perhaps by using a 16-bit dispatch table. We also assume the equipment is provided with a list of standard IDs that it knows how to interpret, as well as a much shorter list of ID length fields corresponding to these standards (possible values of K, defined below).

1. Dispatch on the first byte (256 options); return with 4 integers:

   I = length of length field (0+, 0−, 1, 2, 4, or 6 bytes)

   J = block length B bytes if I = 0 (2, 4, or 6 bytes)

   K = length of ID field in tail (1, 2, 4, 8, or 32 bytes)

   L = presence of descriptor (0 or 1)

2. If I = 0, set block length $B = J$

3. If I = 0−, dispatch (Table 1: not presented here) on M, bits 5 to 8 of the header

   Interpret resulting message and go to 12, or skip directly to 12 if message unknown.

4. If I = 0+, dispatch (Table 2: not presented here) on bits 5 to 8 of header.

   Interpret resulting message and go to 12, or skip directly to 12 if message unknown.

5. If I = 0, read I bytes (yielding block length B), starting at bit 17 of header.

6. If ID field length K not on list of known ID field lengths, go to 12.

7. Read ID field of length K bytes, starting at bit 17 − 8I.

8. If ID is not on list of known IDs, then go to 12.

9. If $L = 1$, read descriptor length D (part of descriptor decoding algorithm, not described here).

10. If $L = 0$, then $D = 0$.

11. Go to algorithm specified by ID and execute over a block of B bytes, starting at the end of the descriptor at $I + K + D$ bytes.

12. Jump to end of block (B bytes long) and read next header.

5.3 Descriptor Specification

5.3.1 Introduction

Publicly readable descriptors may or may not be incorporated in any data block, as indicated by one of the bits in the header key. They would convey auxiliary information concerning the nature of the associated data, such as authorship, distributorship, ownership, intellectual property restrictions, sampling patterns, filtering employed, color, nonlinearities, etc. This information would generally be in the form of identification numbers assigned by standards bodies, although options for conveying text, programs, or other data would exist. Like the header, descriptors would indicate their length so that equipment could skip past if desired, and they would have optional provisions for error protection. An efficient ASN.1 equivalent for the descriptor definition proposed below could also be developed.

5.3.2 Descriptor Architecture

The descriptor is divided into three parts: a 2-byte "key," a "core" ranging from 0 to 8 bytes, and the "tail" of length defined by the core. The key unlocks the core, which defines the length of the descriptor, an indication of the nature of the contents of the descriptor, and the nature of any optional error protection for the core. The tail consists of a series of descriptor identification numbers, similar in concept to the standard identification numbers provided in the header. For each camera type, nonlinear luminance mapping, movie producer, filtering algorithm, royalty payment procedure, etc., there could be a separate descriptor number assigned or registered by appropriate standards bodies, indicated in a manner also similar to that of the header. To properly convey this list of descriptor identification numbers, the tail also contains fields giving the number of such descriptor standards, the length and type of each such standard number, and the nature of any optional error protection employed. The tail also supports delivery of fields of text in any of a large number of languages, such as English or Portuguese, as well as computer languages such as C, Postscript, etc. Such software elements would permit the decoding procedures to be defined explicitly, if desired.

A potential area for future improvement is development of a more universal subset of descriptor elements for widespread usage. It would include parameters such as resolution, raster definition, bit packing, etc.

5.3.3 Descriptor Key Definition

5.3.3.1 Descriptor Key Architecture

The descriptor key consists of three parts: (1) a 4-bit "type" field that characterizes the contents of the descriptor, (2) a 4-bit "length type" field that indicates the format of the descriptor length field, and (3) an 8-bit "protection" field for the key.

5.3.3.2 Descriptor Key: 4-Bit "Type" T Field

The purpose of the T field is to indicate: (1) whether or not this descriptor contains a public index (1 bit); (2) whether the descriptor length field in the core (if any) is 2 or 4 bytes long in integer format (1 bit): and (3) which of four error-protection options are being employed (2 bits). The four descriptor error-protection options are: (1) no protection, (2) protected core plus unprotected tail, (3) both core and tail protected, and (4) both core and tail doubly protected.

5.3.3.3 Descriptor Key: 4-Bit "Length-Type" DLT Field

This field contains the length of the descriptor after the core, in bytes, unless its contents are "zero, zero, zero, zero" (for descriptor tails longer than 16 bytes), in which case the length is given by the core in a field which is either 2 or 4 bytes long, as specified in the T field (see 5.3.3.2).

5.2.3.4 Descriptor Key: 8-Bit "Protection" P Field

The function of the P field is identical to that of the 8-bit error-protection field of the header key; it protects the 2-byte descriptor key only.

5.3.4 Descriptor Core Definition

5.3.4.1 Descriptor Core Architecture

The descriptor core consists of three parts: (1) a field defining the descriptor length (0 to 4 bytes); (2) a field indicating the nature of the contents of the descriptor (0 or 2 bytes); and (3) an optional protection field for the core only (0, 1, or 2 bytes). The total length of the descriptor core thus ranges between 0 and 8 bytes.

5.3.4.2 Descriptor Length Field

This field is of length zero if the descriptor length specification has been preempted by the length field in the descriptor key; otherwise it is either 2 or 4 bytes long in integer format, as determined by one of the bits in the T field of the descriptor key (see 5.3.3.2). The length of the descriptor field as presented in the core is defined as including all the bytes in the descriptor, including those in the key, core, and tail.

5.3.4.3 Descriptor Core Contents Index

This field contains 0 or 2 bytes, as indicated by one of the bits in the descriptor key T field (see 5.3.3.2). In its 2-byte form it indicates whether or not the following descriptor contains information concerning any of 16 categories of information about the data stream. Among others, these categories of information include synchronization reinforcement, error-protection data, encryption keys, packet priorities, authorship, distributorship, time or date of any event, ownership, intellectual property restrictions, sampling patterns, filtering history, color, nonlinear mappings employed, etc. The last bit of the index is zero if this descriptor is the same as the previous one associated with the same header ID. The purpose of this contents index is to spare equipment the burden of decoding descriptors when their contents may be of no interest. This is particularly so when a long sequence of descriptor elements is repeated periodically to aid certain users having only segments of the data stream available to them. Users of longer segments could therefore ignore such data more readily.

5.3.4.4 Descriptor Core Parity Protection

This field would contain 0, 1, or 2 bytes, as indicated by two of the bits in the descriptor key T field (see 5.3.3.2). These bytes would protect the core only, using codes similar to those employed in the header.

5.3.5 Descriptor Tail Definition

5.3.5.1 Descriptor Tail Architecture

The descriptor tail consists of four fields:

1. The element-number field, which indicates the number of independent descriptor identification numbers contained in this descriptor; its length ranges from 4 bits to a maximum of 2 bytes.

2. The descriptor element length-type field, which specifies the lengths of each of the descriptor identification numbers contained in the following field, together with their respective types; these types include identification number types similar to those employed for indicating standard numbers in the header, as well as supporting transmittance of text and computer programs; its length is typically 2 to 4 bits per descriptor identification number.

3. The descriptor identification number field, which consists typically of one or more descriptor standard numbers, each in a 1 to 6-byte format or appearing as a sequence of text or code; the total length of this field approximates several bytes per descriptor element, or substantially more if text or code is incorporated.

4. The protection field, as defined jointly by the protection option indicated in 2 of the T bits (see 5.3.3.2) in the descriptor key combined with the indicated descriptor length: longer descriptor lengths would require more bytes of protection for any indicated level of protection.

5.3.5.2 Descriptor Tail Element-Number Field

This fields consists of one, two, three, or four 4-bit words indicating the number of descriptor elements contained within this descriptor. Each 4-bit word contains 3 bits indicating the number of elements, and 1 bit indicating whether or not an additional 4-bit word is appended, up to a maximum of four words. Thus one 4-bit word will suffice for 0 to 7 descriptor types, which normally should be sufficient. Two concatenated 4-bit words offer up to 2 exp 6 = 64 possible elements. Three words can accommodate up to 512 elements, while use of all four words (2 bytes) can accommodate more 8000 elements, which should be sufficient and is the maximum number per header allowed under this protocol.

5.3.5.3 Descriptor Tail Element Length-Type Field

This field consists of a series of extensible 2-bit words, one sequence of such words applying to each descriptor element. In most cases a single 2-bit word would suffice: the options here are that a 1, 2, or 3-byte field is reserved for the associated descriptor identification number; the fourth option is that two 2-bit words are being employed. If the second 2-bit word is employed, the associated options are that 4, 5, or 6 bytes are being employed to indicate the associated descriptor identification number; this accommodates up to 10 exp 12 possible identification numbers, which should be adequate. The fourth option available for the second 2-bit word indicates that an additional 4-bit word is to be interpreted. This 4-bit word offers 16 additional options, the first of which is that following the 4-bit word, a 1-byte word specifies the length of the descriptor identification field.

The remaining 15 options are of similar form, but indicate that types of descriptor data are being employed other than the standard identification number type. For example, type 2 would indicate that ASCII text was being employed in a language indicated by the first character of the text stream; thus 256 possible languages can be used. Types 3 to 16 would indicate which of several possible computer languages or image description formats were being employed, such as C, Postscript, etc. If it is felt that the 15 possible languages available under the 4-bit extension option in the element length-type field is inadequate, then additional 4-bit fields could be appended by using an extension bit, or by using 1 bit in each 4-bit field to indicate additional 4-bit fields are appended and can be interpreted as were the series of descriptor tail element-number 4-bit words. (Alternatively, the 4-bit word could be reduced to 1 or 2, with the understanding that the language is specified by the first following 2 bytes.) Definition of these options is left to the next step in the SMPTE standards definition process.

5.3.5.4 Descriptor Tail Standard Identification Numbers

If one or two 2-bit words have been previously employed to indicate the length of the descriptor identification number, than 1 to 256 bytes may be employed for the ID number itself. The formats for each of these options are indicated below.

1 byte	International standard established by single designated authority
2 bytes	16-bit standard number designated by authorized international standards body (bodies)
3 to 256 bytes	1 byte indicating the sovereign state, and the remainder (2 to 255 bytes) being available for the standards number

The sovereign state and standards numbers would be designated using procedures similar to that specified in the header. Note that the longer standards numbers permit subdesignations under the sovereign state indicator for subsidiary standards bodies, including individual corporations, institutions, and even individuals. The longer fields also permit use of user-defined bits that can be assigned at execution, thus providing a data field. Such data fields could be used for conveying dynamically changing information such as average luminance, audio gain, etc.

5.3.6 Descriptor Tail Error Protection

This field could be concentrated at the end of the descriptor or distributed throughout to simplify processing. The number of bytes and protocol employed for this purpose would be determined uniquely by the 2 bits in the descriptor key T field and the descriptor length, as specified in the descriptor core descriptor length field or in the descriptor key length field. Definition of these protection strategies might parallel those employed in the header and remain to be defined more fully.

5.4 Transport Header/Descriptors

5.4.1 Motivation and Objectives

Section 5.2.7 discusses several reasons why providing error protection, synchronization reinforcement, packet priority, an higher level encryption to header/descrip tors could pose problems for interpretiv hardware if the data is excessively nois Although the solutions suggested ther will accommodate most error environ ments, still more serious situations can b handled using a transport header/descrip tor block such as described here. Such transport block has the additional advan tage that if insufficiently protected data i moving into a more hostile transport envi ronment, additional protection can be in corporated in the transport block withou having to redefine the input blocks. Simi larly, such transport blocks can be re moved without penalty when moving int more error-free or otherwise benign envi ronments.

The objective here is to suggest how th architecture of such transport blocks ar consistent with the header/descriptor defi nitions presented above, but not to defin all the details. Thus establishment of th header/descriptor standard can procee without waiting for all details of the tran port block to be resolved. It would be use ful to resolve such details, however, befor users of the standard adopt inferior meth ods of addressing the same transport prob lems.

The principal motivation for definin transport blocks is to avoid proliferation o standard-specific alternatives for addres ing such transport problems. Such prolif eration could increase the cost of decodin equipment that would have to accommo date all these possibilities. In a high-erro environment, executing the multipl search strategies necessary when synchro nization is lost or heavy errors exist, could become prohibitive. For this reason it i important to have only a few standard op tions for certain aspects of the transpor block. Defining an efficient small set i beyond the scope of the present effort and would take considerable study. Therefore it is reasonable to assume that this study would be completed after any initial head er/descriptor standard is specified. One il lustrative candidate for such a set appears here in Appendix A; it is intended only to initiate discussion of these issues.

5.4.2 Architecture of Transport Blocks

Transport header/descriptor blocks would consist of a single header/descriptor, where the header would specify an international standard ID indicating which type of transport block was involved; only a few such types would ever be defined. The descriptor of the transport block would convey up to eight different elements:

1. Descriptor table of contents (standard format defined earlier).
2. Synchronization reinforcement bits, not error protected.
3. Error-protection bits for the transport header and its attached following

header/descriptor.

4. Encryption key for deciphering the descriptor, if any, in the following block.

5. Block priority: determined by data originator, indicating relative priority of the following block concerning interpretation or transmission in cases where inadequate capacity is available. Authorization keys may also be needed to verify priority in certain cases. Price bidding could be supported here too.

6. Authorizations and fee mechanisms for alteration or use of data.

7. Block sequence numbering and timing-reconstruction information.

8. Padding to yield one of the very few allowed lengths for the transport block corresponding to the international header standard ID.

In addition to these elements there would also be the traditional field in the descriptor defining its length, although interpretation of this length would be unnecessary because the transport lock length is specified by the header, and there is no data payload following the descriptor.

The six main descriptor elements would convey information using traditional descriptor standard numbers, where long numbers accommodating an adequate number of user bits could be employed. Defining these descriptor standard numbers is the task that can and probably must be postponed until the technical tradeoffs associated with different choices are better understood. Thus standards for headers and descriptors can and should be adopted in advance of these descriptor definitions for transport blocks.

5.4.3 Decoding Issues for Transport Headers

Decoding such a transport header in a high-error environment would be relatively straightforward. First, if synchronization had been lost, synchronization would be established. Initially this might be done assuming a low-error environment. If the environment is noisy, then each of a few possible synchronization blocks would be sought, including periodic repetition of legal header keys, assuming that key bursts might have been employed for this purpose. Because all possible synchronization words might be sought, it is important to have only a few legal ones if they are many bytes long. Because the sync reinforcement bits could be only in a small number of positions relative to the beginning of the transport block, each such position could be tested for consistency with the associated error-protection bits, which are also located in a small number of possible locations. Confirmation of synchronization follows if legal headers come immediately after the indicated end of any block. Once the transport block is synchronized and error corrected, the remaining descriptor fields containing any encryption key for the descriptor in the following block, or any packet priority information can be deciphered.

One principal new constraint should be imposed by the standard on manufacturers of equipment handling this header/descriptor standard. If transport blocks are to be useful, such equipment must never insert data between a transport block and the following block to which it applies. Transport systems should try not to scramble the sequence of data blocks in any event, but if a transport block should accidentally be prepended to the wrong data block, the packet priority, encryption key for the data block descriptor, and the error protection for the header/descriptor could be inappropriate, resulting in a scrambling of the interpreted header/descriptor. Such scrambling would also typically cause local loss of synchronization, particularly in high-error environments. Since transport blocks would normally be quite brief compared to typical data blocks, such a constraint should not be difficult to satisfy. Such transport blocks could also be added or subtracted at will by a given transport layer, however, provided they are appropriate to the blocks which they precede.

If a data stream is entered in the middle of a transport block, then confusion might result. To protect against this unlikely possibility, equipment might choose to wait until the second valid header is intercepted before commencing decoding.

6. Abstract Syntax Notation 1 (ASN.1) Header/Descriptor Architecture

6.1 Background

Abstract Syntax Notation 1 (ASN.1) is an existing ISO/CCITT Standard in common use within the computer and telecommunications industries. Within the ASN.1 framework, it is straightforward to define an SMPTE header/descriptor that meets the objectives described in Sections 1 to 5 above, and it would leverage existing tools, expertise, and administrative structures.

ASN.1 is derived from earlier work at Xerox PARC on Courier (late 1970s). An early version of the notation (c. 1984) was used in the first draft of the CCITT X.400 series of recommendations on message handling systems (i.e., electronic mail). ISO and CCITT then jointly developed ASN.1 for use within the OSI presentation layer (c. 1988).

ASN.1 is now widely used in a range of standards activities, including the CCITT X.500 directory service and both the OSI and Internet network management systems. Over the years, a collection of software tools and utilities to support ASN.1 has been (and is being) developed.

6.2 Concepts

ASN.1 is an extensible notation for describing data that is to be exchanged by transmission or storage. It is much like a programming language, such as C and Pascal. There are several simple types, such as integer, real, and octet string (i.e., byte string), and constructor types that can be used to build arbitrarily complex data structures, including hierarchical representations (e.g., packet within packet).

An ASN.1 header can be thought of as an envelope that contains, for example, a single video frame. ASN.1 supports the notion of embedding, which allows one or more data structures to be contained within another. Thus, a sequence of frames can be embedded within an outer header (or envelope) that labels a program segment. This can be taken to coarser granularity, e.g., shots, scenes, programs, etc. Similarly, it can be taken to finer granularity to embed audio tracks, closed captioning, descriptors, etc., within individual frames.

A key feature in ASN.1 is the separation of how the data is described (Abstract Syntax) and how data is encoded (Basic Encoding Rules, or BER). Data structures are described in a human-readable syntax and automatically translated into the bits and bytes for transfer. When a new data structure (or type) is defined, its representation is automatically generated. Furthermore, deployed ASN.1 compliant systems will be able to interpret new structures without hardware modification.

The following summary description of ASN.1 presents only enough detail to motivate its use for the specific needs of a header/descriptor. For formal definition of ASN.1, refer to ISO 8824/8825 and/or CCITT X.208/209. A more accessible description can be found in: Marshall T. Rose, *The Open Book: A Practical Perspective on OSI*, Prentice-Hall, 1990.

6.2 Basic Encoding Rules (BER)

All ASN.1 types, whether a simple type or a structured type, can be encoded using the same basic format of three fields:

[tag] [length] [value]

The three fields together make up a data item. Each field is variable in size to accommodate arbitrarily complex substructures and encodings. A simple type, such as an integer, requires only a few bytes. A structured type, such as a long byte string, can be Mbytes, Gbytes, or larger as necessary to contain the payload data value. The basic format is inherently self-identifying and extensible.

A data stream is a sequence of items, each of which can be structured or nested. Thereby, one can define arbitrarily structured data for both header and descriptor, including nested packet-within-packet structures.

"Tag Field". The tag field specifies the type of the item value. Several simple and structured types (integers, character strings, etc.) are universally defined in the ASN.1 standard, and are recognized in all

compliant environments. Also, one can define types that exist within the specific environment of an application or communications context. A tag is principally encoded as a single byte, but can be extended. In the BER, the tag is encoded as:

<2> <1> <5> —bits/field
[class | p | tag number]

The tag field allows a receiver to "parse" the incoming data stream, selecting those components/types in which it is interested and bypassing others.

The EXTERNAL type is of particular significance. In essence, it is a universal header for the data that it encapsulates. The EXTERNAL type is described further below. Other types are relevant to use in a descriptor.

"Length Field". The length field indicates the size of the value. It is an integer of one or more bytes that specifies the number of bytes in the value field. In the BER, the short form of length is encoded as a single byte, and can indicate lengths of 0..126 bytes of value. In extended form, the first byte specifies the number of bytes of length. Length is encoded as:

short form : 1 byte : [0bbbbbbb] : lengths of 0..1126 bytes

extended form : n bytes : [1nnnnnnn] [bbbbbbbb] [bbbbbbbb] ...n

Only the number of bytes needed to specify length are used. Thus, the length field is compact. The extensibility of the length field permits a maximum length field that is 126 bytes to specify a length of ~2^1008 or ~10^303.

Note that a common length specification is used regardless of the associated data item. Accordingly, there is no need to invent custom length encoding schemes for each new data item.

"Value Field". The value field is the value in the type specified by the tag. It is a string of bytes the number of which are specified by the length field. It is encoded as defined in the BER for that type.

6.4 ASN.1 EXTERNAL Type (Universal Header)

An ASN.1 EXTERNAL type is a universal header. All ASN.1 compliant protocol interpreters can extract and interpret an EXTERNAL without ambiguity. The definition of EXTERNAL is quite flexible, but that flexibility is not needed here to meet the basic objectives of a universal header.

A simplified EXTERNAL type is encoded as:

[tag = class = 0,p = 1,number = 8] [length] [object id] [payload]

"Tag" and *"length"* fields are encoded as described above. *"Object ID"* provides unambiguous self-identification for the header. *"Payload"* is a sequence of bytes that are interpreted according to the standard indicated by the object ID.

"Object ID" is itself an ASN.1 type with tag, length, and value fields, and is encoded as:

[tag = 0.0.6] [length] [id value]

"Object ID" value is a sequence of bytes that represent the hierarchical identifier for the referenced standard. ID values are assigned, registered, and administered by CCITT and ISO in the course of standards development. Or, ID assignment can be delegated to member bodies or companies or organizations (thereby, SMPTE could assume responsibility for administering a portion of the ID space.) The root prefix values are:

CCITT [0]		
recommendation	:	CCITT committees
question[1]	:	CCITT Study Groups
administration[2]	:	country Ptts (country code)
network operator	:	X121 organizations
ISO[1]		
standard[0]	:	ISO standards
registration authority[1]	:	ISO authorities
member body[2]	:	member bodies (country code)
identified organization[3]	:	organizations
joint ISO CCITT[2]	:	assignment delegated to ANSI

A few prefixes are of particular note. iso.standard registers all ISO standards. ccitt.administration and iso.memberbody are assigned to sovereign bodies (identified by their international telephone country code). iso.organization is assigned to international organizations. These cover virtually all the situations under which a header identifier will need to be assigned.

An ID value is encoded as a sequence of bytes. The first two levels are encoded in the first byte — $a.b. = 40^*a+b, a <= 3, b <= 39$. Remaining levels are encoded as one or more bytes as needed to represent the numerical value for that level. If a value is greater than 127, it requires more than one byte; the MSB of the byte is set to indicate that the value is continued in the next byte, and so on. For example, iso.standard.jpeg is 3 bytes:

iso.standard.jpeg :: 1.0.10918*
= [40] [128+85] [38]

Similarly, Group 3 Fax is identified as:

ccitt.recommendation.t.4 :: 0.0.20.4
= [0] [20] [4]

The extensibility of the ID value field permits up to 126 byte long IDs to specify distinct IDs numbering to ~2^882 or ~10^265.

"Payload" is an ASN.1 type with tag, length, and value fields, encoded as a sequence of bytes that are interpreted according to the standard indicated by the *"Object ID"*.

[tag = 2.0.1] [length] [payload value]

6.5 ASN.1 Descriptor

In addition to the EXTERNAL type (for use as header), ASN.1 defines other basic types to represent a variety of values, including: booleans, integers, reals, byte strings, character strings, universal time code, etc., and constructions of values into arbitrary data structures.

Although a distinction is drawn between header and descriptor in this report, the ASN.1 approach permits a single mechanism to serve both functions. The full benefits of ASN.1 become apparent when it is applied to the descriptor. Especially the ability to construct new types and to incorporate references to other standards. (See current ISO work on Image Interchange Format [IIF] for an example of ASN.1 use in defining descriptors.)

6.5 How ASN.1 Addresses Objectives

The following describes how an ASN.1 header/descriptor addresses the objectives stated at the beginning of this report.

Universality — ASN.1 header/descriptor promotes and enhances universality:

• It complies with and recognizes existing standards and practices. All existing and future ISO/CCITT standards are uniquely identified by an ASN.1 Object ID.

• It addresses the issues of sovereignty. ISO/CCITT administers and delegates assignment of Object IDs to subcommittees, member bodies (by country code), and organizations. The complexity of this task and the benefits of leveraging existing administrative structures should not be underestimated.

• It facilitates coordination among television, telecommunications, and computer industries.

• It sets a minimal level of compliance for low-cost receivers. Furthermore, the advanced stage of definition, tools, and expertise will facilitate the rapid deployment of header-compliant devices.

Longevity — ASN.1 has inherent longevity:

• All fields of ASN.1 types (tag, length, value) can be extended. Payload lengths from a few bytes to ~10^303 can be represented. Similarly, Object IDs can range from a few bytes to ~10^265 bytes.

• It has a preexisting registry and is self-maintaining. ISO/CCITT already registers, administers, and delegates assignment of Object IDs.

• It defines immutable identifiers. Once an Object ID is assigned, it exists "for all time." In the future, when an old

*At this writing, JPEG is nearing but not yet an ISO standard. Thus, though thought to be correct, the number here (1.0.10918) is not yet official.

ASN.1 header is recognized, there is no ambiguity to the referenced standard.

Extensibility — ASN.1 is inherently extensible:

• All fields of ASN.1 types can be extended without redefinition.

• New types can be defined and their encoding automatically generated without the need to introduce new rules.

• Since ASN.1 is fully defined, any compliant receiver and equipment are guaranteed to be able to recognize future extended ASN.1 headers.

Interoperability — ASN.1 is inherently interoperable:

• It has a well-formed public definition.

• It is already in use in several important applications and industries.

• It complies with existing and developing standards (including image standards).

• It allows standards and structures to cross-reference each other. The video data stream can contain structures defined by other standards, and vice versa.

• It permits the same information to be interpreted in different ways within different domains without prejudice. To the video industry, the information is a continuous video stream. To the telecommunications industry, the same information is a sequence of bits to be transmitted. To the computer industry, that sequence of bits is interpreted as data structures.

• It permits independent formal definition of intellectual property protection, encryption, and other source coding descriptors by appropriately sanctioned expert groups (either ad hoc or extant).

• Experts groups and standards bodies can autonomously develop and evolve specifications within the domains of their expertise. The ASN.1 cross-referencing capability achieves a degree of interoperation and coordination among parallel activities — coding details are automatically resolved, avoiding redundant and/or conflicting efforts.

Cost/performance effectiveness:

• It is straightforward to recognize and decode with a single uniform procedure.

• Uniform decode hardware and software can be shared among industries and applications yielding economies of scale.

• Existing tools and expertise can reduce the time and cost of deployment.

Compactness:

• ASN.1 is compact, but not so much as to complicate decoding or to compromise other objectives. A typical ASN.1 header would be ~15 bytes for a 1-Mbyte payload (i.e., ~.0014% overhead).

• A 7-byte ASN.1 header can be realized using the standard's indirect reference option. (In fact, a 2-byte context specific header may be realizable.) Given typical payload sizes, however, it is unlikely that this level of compactness will be required — small payloads are either aggregated or infrequently used, especially in the video domain.

Rapid Capture:

• An ASN.1 header is signified by its first byte tag field (equal to 8) followed by length, object ID, and payload. Thus, an ASN.1 header is straightforward to recognize in the data stream.

• ASN.1, like any length/identifier header, provides early identification of the payload.

Editability:

• ASN.1's structuring capabilities permit arbitrary editing, sequencing, structuring, and embedding of payload streams. All of this is accomplished within a single uniform mechanism, and without requiring unnecessary decoding of the payload itself.

Appendix A — Transport Header

A.0 Introduction

The header-descriptor design provides for binding of a transport header to the main header. This is accomplished only by mandating that the transport header not be separated from the main header and its associated data, which it is transporting. The function of the main header is to identify the data that follows. The function of the transport header is to help the main header and possibly its payload in its journey from origination to destination.

The following transport header design is "work in progress," and therefore is meant to be an example rather than the final structure. It supplements the discussion in Section 5.4. The design issues and interactions are a bit complex, so the principles of the transport header design are best illustrated by way of a correctly designed example. If adjustments are made to this design, then care must be taken concerning effects of such adjustments to the rest of the design and the functioning of the transport header. In particular, there is a rigid requirement that certain fields be prespecified as to length, or as to length-specification (type fields) in their meaning.

It is necessary that the transport header be totally independent of any standards described by the main header. This is required because it may be necessary to change transport characteristics for all main headers on a given data stream, irrespective of the standards or formats of those main headers.

An example might be the need to provide improved error protection when moving a data stream from a high-reliability fiber to a high-error-rate radio frequency transmission.

A.0.1 ASN.1 Transport Header Yet to be Designed

The transport header design example illustrated here works together with the header/descriptor design, as described in Sections 1 through 5 of this SMPTE report.

No transport header has yet been designed for the ASN.1 syntax, illustrated in Section 6. Thus, the ASN.1 syntax is, at present, only suitable for error-free channels that preserve the data and its order without contention from source to destination. In order for the ASN.1 syntax to meet its ojectives of interoperability with imperfect or congested channels or media, a mechanism similar to the transport header example shown here will be required. It is hoped that a transport header design might be developed for the ASN.1 syntax system, possibly modeled on the transport header illustrated here, together with the main header/descriptor design. The enormous flexibility of ASN.1 syntax must be tempered to provide a limited number of options for transport headers, each with appropriate protection/correction mechanisms. It is hoped that registration rules and flexibility in ASN.1 can also be used to provide a suitable format transport header design that is properly restricted. Byte alignment is also part of the structure of length fields. If bit alignment is needed for ASN.1, then further suitable adjustments will be required.

Another method to provide transport assistance is to convert from the main header/descriptor design of Sections 1 through 5, to and from the ASN.1 notation. Since a transport header design is available for the main header/descriptor, conversion to this header would provide access to a transport header. When reentering error and contention-free environments, the header/descriptor could be reconverted to ASN.1 syntax.

If the ASN.1 method becomes popular, then it is hoped that a suitable ASN.1 transport mechanism might be developed.

A.1 Design Objectives for the Transport Header

• The transport header must be standard-ID independent, so that it can apply to all header standards equally and uniformly across the entire data stream.

• The transport header should be removable without damage to the function of the main header that it is helping to transport.

• A transport header must be able to be added to any header format or descriptor format without changing any of the meaning.

• The transport header formats should support "in the clear" protection of the main header and payload, where the bits are not altered, so that the transport header can be added and removed without any adjustment of bits within the main header, its descriptor, and its payload.

• In addition to support for "in the clear" protection, more

efficient protection should be supported (such as Reed-Solomon), where the header, and possibly its payload, are encoded.

- The transport header architecture should support one or more mechanisms for correcting burst errors.
- Optional support should be provided for encryption of the main header's descriptor, as well as the data payload.
- The transport header should support authorization and use information, in helping the transport system determine which destinations are appropriate for further or final distribution.
- The transport header architecture should support backward play through the data stream.
- The transport header should support optional rapid header synchronization capture.
- The transport header should support communications networks by providing information concerning the data's priority and value.
- The transport header should support timing reconstruction when utilizing networks that distort timing or ordering of data.
- The transport header should support data ordering requirements, when utilizing networks which might reorder the data.
- The transport header architecture should strike a balance between the opposing forces of flexibility and ease of use. Thus, a small number of options should be carefully chosen for maximum flexibility, with the small number of options allowing a simple interpretation. It is the fact of having a small number of options that allows easy interpretation.
- For transport functions that encode the header and/or its data, a simple "in the clear" length field should allow devices that cannot decode such data to skip to the next transport header.
- For devices that cannot process the transport header, a simple "in the clear" length field should allow such devices to skip directly to the header. Further, such devices should be able to easily interpret the transport header format so they can quickly determine whether the main header is "in the clear," and therefore directly readable.

The following example design meets these objectives.

A.2 Three Types of Transport Header Allocated in the Header Key

In order to meet these objectives, there are three types of transport headers. The first is the basic transport header, which provides the majority of transport capabilities. The second is the redundancy transport header, which is used to protect against burst errors. The third is the reverse transport header, which is used for reverse play.

Three of the 2-byte header key's 256 possible codes are required for the transport headers. In the current 2-byte header key design, 1 byte is dedicated to error protection. The other byte is split into two 4-bit fields.

The first 4-bit field is the "length type" field. Codes 15 and 16 (numbers 14 and 15) are unallocated. These two codes enable the second 4-bit ID field to be used for special purposes such as designating the three transport header types. Thirty-two such codes are available, leaving 29 codes unallocated if 3 codes are assigned to the three transport header types. The basic transport header provides most of the capabilities required.

A.3 Functions of the Transport Header

The functions of the transport header are as follows:

1. Sync reinforcement for those data transport media where it is desirable to rapidly or simply sync to the headers or header-data combinations on switching between streams.

2. Improved error protection via extra protection bits for both the transport header and the main header and its descriptor.

3. Conveyance of priority for the main header and its data, for those cases where a channel may be operating at capacity and thus where channel controllers must decide which headers and their payloads it must drop. Authorization keys may also be needed in order to verify priority. Network accessing methods, such as pricing-bidding techniques, would also be supported here.

4. Encryption and security information for the main header's descriptor, possibly combined with the descriptor's own encryption and security information, in order to protect the data stream following the main header. The protection optionally provided by the main header's descriptor may need to be augmented when the data stream is sent through public or vulnerable exposed channels.

5. Authorization information. Such information would indicate who could receive, who could edit and reassemble with other material, etc. Also, copyright and royalty-fee information would be enabled here. Perhaps automated mechanisms of fee for usage would be supported through this field.

6. Sequence numbers may be added where networks are used for transport, which may reorder packets. In addition to sequence numbers, the combination of the transport header may require information from the main header's descriptor in order for the network to be able to guarantee delivery within known latencies and time windows. For out-of-order delivery, some networks can control the amount of time between a given delivery and the delivery of neighboring data. In the case of images and audio, some devices can accept out-of-order information in their buffers or processing units, but the time is constrained to within one or more frames, or fractions thereof.

7. Timing reconstruction. For those applications where exact timing relationships must be reconstructed from a mixed data stream, the transport header and the main header's descriptor would communicate to provide timing reconstruction information.

8. Reserved for future use.

9. Pad. There is a pad field at the end of the transport header in order to make the length of the transport header plus the main header, its descriptor and, optionally, its payload, come out to a length appropriate for the error-correction protection formats supported in item 2.

The construction of the transport header involves a strict ordering of fields as above, so that the sync reinforcement is always first, the error protection is always second, etc. In this way, if each field is present, its location is easily determined. The error-protection field will always be in a known location, and a "scope of protection" within this field defines those fields that are protected in both the transport header as well as in the main header and its descriptor.

A.4 Redundancy Transport Header

The redundancy transport header provides a special function for error protection against burst errors on the data stream. Improved error protection against burst errors is achieved via redundant copies of future and previous headers (these transport headers providing redundancy will not be bound to a main header, and therefore represent an exception to the binding property of transport headers).

Such special redundancy transport headers will "stand alone" and be occasionally interspersed in the data stream. They will contain error-protected copies (via either the efficient or inefficient method) of some future or previous header, together with a pointer to that header, and a number indicating how many headers forward or backward will be traversed before reaching that header. The previous copies are useful for going backward through a data stream, and for reconstructing damaged data streams on physical media, such as disk.

Using a separate header key code, the redundancy transport header has the format shown in Fig. A1.

The transport header key contains a separate special code indicating that this is a redundancy transport header containing a copy of a future or previous header. This field could possibly be followed by a length field, indicating how to skip past this duplicate header.

The pointer to the header being duplicated is analogous to a length field, but it points past several headers to the header being duplicated. It is a signed number so that it can reference previous as well as future headers. Thirty-two bits of protection are pro-

vided. A number of headers forward or backward is provided, indicating the location of the header being duplicated in number of headers rather than via a pointer (length).

The maximum and minimum millisecond tolerance fields indicate the tolerances for separation times between this copy of a previous or future header and the header itself. Information about separation constraints is provided by these fields for channels that reorder, insert, and remove data.

A copy of the transport header is provided, if there is a transport header on the header being copied.

Transport Header Key Unique Code	Protection	Pointer To Header Being Duplicated (Signed Number)	Protection	
	<-		<-.......	more ->
8	8	32	32	

Number of Headers Forward (or Backward) (Signed Number)	Protection	Maximum Millisecond Tolerance From Header	Minimum msec Tolerance From Header	
	<-			more ->
16	16	8	8	

Protection for max and min tolerances	Copy of Transport Header (if present)	Copy of Main Header/Descriptor
16	(length varies)	(length varies)

Figure A1. Format of redundancy transport header.

Transport Header	Main Header/ Descriptor	Data Payload	Reverse Transport Header

Figure A2. Situation of transport header.

Reverse Transport Header Key Unique Code	Protection	Length Backward To Main Header	Protection	Length Backward To Transport Header	Protection
	<-		<-		<-
8	8	32	32	32	32

Figure A3. Reverse transport header format.

Finally, there is a copy of the complete main header and its associated optional descriptor. None of the payload is duplicated.

A.5 Reverse Transport Header

In order to support reverse reading of the data stream, a reverse transport header can be utilized. The reverse transport header immediately follows the main header. Thus the transport header is situated as shown in Fig. A2.

If the main header is preceded by a transport header, then the reverse transport header will point back to both the main header and the transport header. If the transport header is absent, then the reverse transport header will point back only to the main header, and the length backward to the transport header will be zero.

The reverse transport header format is shown in Fig. A3. This design allows these reverse transport headers to be appended after the main headers to allow backward traversal through the data stream.

It should be noted that when redundant transport headers are in use to protect against burst errors, reverse transport headers must follow each such redundant header. In that case, the length backward to the main header will be zero, and only the length backward to the transport header will have a nonzero value.

Note that this header has a fixed length. Thus, when encountering this header in the forward direction, no pointer to the next header in the form of a length field is required. The fixed length of 18 bytes is predetermined and can be used to skip to the next header. Also, there will never be a payload of data or any attachment to any headers in the forward direction.

A.6 Transport Headers and Device Capture

It is important to remember that all devices that can edit the data stream must preserve the relationship of the preappended transport header to the main header. Also, the postappended reverse transport header must also remain attached to the main header, if present. Thus, when a transport header is read, the following two headers must also be read before assuming the header and its data and transport have been passed.

When capturing a new data stream, it is necessary to read at least two headers to determine if the first main header is valid. If it had been preceded by a transport header that encoded the main header's descriptor and/or payload, then the data will not be readable without the transport header. Thus, capture is not achieved until the second header has been read, which would accomplish the determination of the first valid header and its transport header, if present.

A.7 Transport Header Format

The transport header format is shown in Fig. A4.

A.7.1 Header Code in the Header Key

The transport header begins with a normal header "key," consisting of 1 byte of key information and 1 byte of protection, as for the main header. However, the transport header uses a header key code reserved specifically for the transport header. The transport header tail differs from the main header tail and has the format outlined in Fig. A4.

A.7.2 Length to Next Header after Main Header

The next field is a 32 bit length field, which points to the next header after the main header attached to this transport header. This length may be required to allow skipping past the main header and its data payload. This will be required by those devices that cannot provide error-protection decoding, when the main header is protected using the types of error codes that scramble the main header. This length is protected by 32 additional protection bits to provide reasonably robust bit-error correction, using the type of correction that does not scramble the 32 bits of length (e.g., Hamming code).

A.7.3 Transport Header Length

This field indicates the length of the transport header. It therefore forms the mechanism to skip to the main header, if there is no desire to read any of the transport fields, and if the main header is not scrambled due to error protection, encryption, or other operations from the transport fields. This field is also protected by a 32-bit field, using a nonscrambling error-protection code.

A.7.4 Type Fields

There are eight 4-bit type fields, to allow 16 types for each of the eight fields in the transport header. These fields are (1) sync, (2) error protection and correction, (3) priority and authorization and bidding for priority, (4) authorization for data use, (5) encryption protection, (6) sequence number and out-of-order timing margins, (7) timing reconstruction information, and (8) a final field reserved for future use. The 32 bits of type fields are protected by 32 bits of protection/correction code.

A.7.5 Sync

The first transport operation field is sync reinforcement. Sixteen types of codes are possible, each with a permanently assigned length. These types and lengths must be permanently assigned from the beginning. A possible set of 16 assignments for lengths might be as follows.

Two types of sync codes could be used, with custom-designed unique spectral signatures. Each of the two types of codes could have one of the following lengths:

2, 4, 8, 16, 32, 64, 128, and 256 bytes

This would result in a total of 14 sync reinforcement patterns. Sync type 0 would indicate an absence of the sync reinforcement field. Sync type 15 could represent an additional special sync field, with a specified length.

It is necessary for all sync type lengths to be specified in advance. Although the codes for sync themselves can be specified later, they can only be specified once for each of the 15 valid sync types.

It should be noted that the transport header always begins exactly 26 bytes prior to the first byte of the sync reinforcement field. Once the sync reinforcement field has been located, the transport header and the main header have been located.

A.7.6 Error Protection

None of the fields previously described, which precede the error-correction/protection field, will be protected by this field. However, all of the fields following and including this error-protection/correction field, starting at the first bit, will be part of the error-protection/correction group. The protection will extend through the rest of the transport header and on into the main header and its descriptor, and for some of the formats, into the payload as well.

Since sync reinforcement is used to capture the data stream initially, it is probably not appropriate to error-protect this sync. The sync codes are designed to be found within a stream. If protection is needed for the sync field, then special sync protection can be provided within the ample bytes available for sync codes.

Two types of error protection are supported. One type allows the protected data fields to be read, as is, leaving them "clear" but augmenting them with protection. Examples of this type of code are the Hamming code and some common forms of the Reed-Solomon code. The second type scrambles all of the bits in the coding process. An example of this type of code is a convolutional one such as the Viterbi code. In the case of scrambled protection, the fields being protected will be completely unreadable without decoding. If the header format, the header length field, the descriptor type, and many other crucial fields are protected in this

```
--------------------------------------------------------------------------
| Transport | Protection | Length         | Protection |
|  Header   |            | of Transport   |            |
|   Key     |     <-     | Plus Main      |     <-     |
|  Unique   |            | Header Plus    |            |    more ->
|   Code    |            | Data Payload   |            |
|    8      |     8      |    32          |     32     |
--------------------------------------------------------------------------

--------------------------------------------------------------------------
| Length of | Protection | Sync |  Error          | Priority   |
| Transport |            | Type |Protection/      |    &       |
|  Header   |     <-     |      |Correction..|  Valid Bid  |
|           |            |      |  Type           |   Type     | more ->
|    32     |     32     |  4   |   4             |    4       |
--------------------------------------------------------------------------

-----------------------------------------------------------------------------
| Author- | Encrypt | Sequence | Timing | Reserved | Protection          |
| ization |         |  Number  |        |  Future  | For Previous        |
| Type    |  Type   |  Type    |  Type  |  Type    | 8 4-bit Types       |
|    4    |    4    |    4     |    4   |    4     |     32              |
-----------------------------------------------------------------------------
                                                                  more ->

-----------------------------------------------------------------------------
| Sync           | Protection/      | Priority &      | Authorization   |
| Field          | Correction       | Auth for        | Copyright and   |more->
|(16 possible    |(many possible    | Priority/Bid    |  Use Field      |
| lengths)       |   lengths)       |(16 lengths)     | (16 lengths)    |
-----------------------------------------------------------------------------

--------------------------------------------------------------------------
| Encryption   | Sequence     |  Timing      | Reserved     | Pad       |
|  Field       |  Number      |  Field       |   for        | Field     |
|              |  Field       |              | Future       |(variable  |
|(16 lengths)  |(16 lengths)  |(16 lengths)  |(16 lengths)  | length)   |
--------------------------------------------------------------------------
(end of transport header ->)

Followed by a normal main header:
------------------------------------------------------------------
| Main   | Protect |  Header
| Header |   <-    |  Tail
|  Key   |         |          ->
|   8    |    8    |  (length fields, descriptor codes, etc)
------------------------------------------------------------------
```

Figure A4. Transport header format.

way, those devices that cannot decode the error correction would be unable to either read or skip the header. Thus, the transport header will contain a "length field," protected in the augmentation method rather than the scrambling method, which will allow devices to find this length field and thereby skip the rest of the transport header and the entire main header, its descriptor, and its payload. A second "length of transport header" field will also be present and protected via the "in the clear" augmentation method.

A.7.6.1 Intentional Inflexibility

In order for all of the above mechanisms to operate properly and efficiently, it is necessary to limit the number of formats available in the transport header. Perhaps 8 or 16 types of each of the fields should be provided for, with all types being specified in advance. We are using 16 as our example. Thus, there would be 16 error-correcting coded formats and code lengths, 16 sync reinforcement field formats and lengths, 16 priority fields and lengths, etc., with each field format having a specified length.

Since the error correction transport function is the most difficult with respect to format, a very limited number of field options will be provided under the scope of protection. Also, padding, which might be quite long, will be required at the end of the transport header before the main header in order to make the total length of the transport header plus the main header and its descriptor (and possibly payload) to be a simple-to-correct convenient known length corresponding to one of the 16 protection types.

A.7.6.2 Possible Prespecified Protection Types

The prespecified 16 possible protection types might be as follows:

- Type 0 indicates that no protection/correction field is present.
- Types 1 through 10 protect "in the clear," by adding protection bits without scrambling, as in Hamming codes and some common types of Reed-Solomon coding.
- Types 1 through 5 protect the transport header, the main header, and its descriptor.

The types are as follows:

1. Protect all remaining transport header bits, beginning at the first bit of the error-protection/correction field, all main header bits, and all descriptor bits. Do not protect any payload bits. The protection/correction bits are applied on every group of 64 bits. The total length of all fields being protected, excluding the error-code bits themselves, must be a multiple of 64 bits. This is accomplished by the use of pad bits at the end of the transport header.

This implementation requires that memory be available to store the correction bits for the entire length of transport header, main header, and its descriptor.

For type 1, the total length of the protection code field is the total length over 16, with 4 bits for every 64 (68 bits total on 64 bits of data).

2. Same as 1, but total length over 8, with 8 bits protecting every 64 (72 bits total for 64 bits of data).

3. Same as 1, but total length over 4, with 16 bits protecting every 64 (80 bits total for 64 bits of data).

4. Same as 1, but total length over 3 (half as long as the fields being protected); 32 bits protecting every 64 (96 bits total for 64 bits of data).

5. Same as 1, but total length over 2 (the same length as the fields being protected); 64 bits protecting every 64 (128 bits total for 64 bits of data).

Types 6 through 10 protect the payload in addition to the transport header, the main header, and its descriptor.

6. Same as type 1, except protect the payload as well. For type 6, the total length of the protection code field is the total length over 16, with 4 bits protecting every 64 (68 bits total 64 bits of data).

7. Same as 6, but total length over 8, with 8 bits protecting every 64 (72 bits total for 64 bits of data).

8. Same as 6, but total length over 4, with 16 bits protecting every 64 (80 bits total for 64 bits of data).

9. Same as 6, but total length over 3 (half as long as the fields being protected); 32 bits protecting every 64 (96 bits total for 64 bits of data).

10. Same as 6, but total length over 2 (the same length as the fields being protected); 64 bits protecting every 64 (128 bits total for 64 bits of data).

Types 11 through 15 protect by scrambling, as in convolution coding techniques such as Viterbi coding. Types 11 and 12 protect the transport header, the main header, and its descriptor.

11. Protect all remaining transport header bits, beginning at the first bit of the protection/correction field, all main header bits, and all descriptor bits. Do not protect any payload bits. The protection is applied using 16 bytes on every group of 144 bytes. The total length of all fields being protected, including the error-code bits themselves, must be a multiple of 144 bytes. This is accomplished by the use of pad bits at the end of the transport header.

This implementation requires that memory be available to store the correction bytes for the entire length of transport header, main header, and its descriptor. For type 11, the total length of the protection code is 16 bytes for every 144. Thus, there are 16 extra protection bytes in each 144 bytes being stored, resulting in 128 bytes of data after decoding.

12. Same as 11, but with 16 bytes protecting every 80 bytes, resulting in 64 bytes of data after decoding. The total of all lengths, beginning at the first bit of the protection/correction field, must be a multiple of 80 bytes.

Types 13 and 14 protect the payload in addition to the transport header, the main header, and its descriptor.

13. Same as 11, except the payload is also protected.

14. Same as 12, except the payload is also protected.

15. Same as 11, except 4 bytes protect 32 bytes (total of 36 bytes for 32 bytes of data). This format is for short header formats, and provides no payload protection.

A.7.6.3 Interleaving

In addition to the above mechanisms for error protection, some of the error-protection formats can invoke interleaving. Interleaving can substantially reduce the problems associated with burst errors. Although the initial part of the transport header is subject to being "wiped out" by a burst error, presumably a copy of this section could be available previously in a redundancy transport header. Thus, once the protection format has been determined, then the rest of the transport header, beginning at the error-protection field, plus the main header, its descriptor, if present, and optionally the data payload, can all be protected from burst errors via interleaving in addition to error-correction methods.

Predefined interleaving methods can be incorporated with some of the types discussed above. Because interleaving is likely to involve as wide a spacing as is feasible, there will be a tradeoff between the length of the protected field, and natural multiples of the error-protection sizes. The error-protection group sizes for Hamming-type codes are much smaller than the error-protection group sizes for Reed-Solomon-type codes. Interleaving must be some multiple larger again. Thus, useful interleaving may be restricted to longer lengths of fields being protected. One possibility is to have the interleaving spacing be the error-protection group size divided into the total length.

However, this near-optimal format requires some complexity in unwinding the interleaving. For long protected fields, this may also involve a buffer that is the length of the field. Thus, there are potential issues to investigate with respect to how to universally and generally specify a powerful interleaving technique.

A.7.6.4 Error Detection

There is no provision for simple error detection in the above type examples. Such detection could be provided via cyclical

redundancy code (CRC), fire code, checksum, parity, or other check method. Such may be useful in some cases. However, the focus on error correction is based on the need for headers to be interpreted without error in order to serve their function.

The descriptor in the main header can be used for detection codes for data payloads that should be checked but need not be corrected. This need not be standard-specific, since the descriptor can be standard-independent. Thus, error detection, as opposed to correction, is more appropriate in the descriptor than in the transport header.

A.7.6.5 Parameters of Error Protection

The parameters of error detection shown in the above type examples need further investigation and refinement. The lengths and protection ratios proposed are known to be implementable in existing hardware and are expected to be convenient in practice. However, further investigation of optimal parameters for error protection may help refine or revise the parameters suggested above.

```
Type 0. 1 byte:
--------------
| Priority |
|  8 bits  |
--------------

Type 1, 2 bytes:
--------------
| Priority |
| 16 bits  |
--------------

Type 2, 4 bytes:
--------------------------------
| Priority |  Authorization |
| 16 bits  |    16 bits     |
--------------------------------

Type 3, 8 bytes:
-----------------------------------------------
| Priority | Bid/Value |  Authorization |
| 16 bits  |  16 bits  |    32 bits     |
-----------------------------------------------

Type 4, 16 bytes:
-----------------------------------------------
| Priority | Bid/Value |  Authorization |
| 4 )ibytes|  4 bytes  |    8 bytes     |
-----------------------------------------------

Type 5, 32 bytes:
-----------------------------------------------
| Priority | Bid/Value |  Authorization |
| 4 bytes  |  4 bytes  |    24 bytes    |
-----------------------------------------------

Type 6, 64 bytes:
-----------------------------------------------
| Priority | Bid/Value |  Authorization |
| 8 bytes  |  8 bytes  |   48 bytes )i  |
-----------------------------------------------

Type 7, 128 bytes:
-----------------------------------------------
| Priority | Bid/Value |  Authorization |
| 8 bytes  |  8 bytes  |   112 bytes    |
-----------------------------------------------

etc.
```

Figure A5. Field formats.

A.7.7 Priority and Authorization or Bid for Priority

Finite bandwidth resources, such as satellite channels, long fiber channels, long real-time computer channels, terrestrial broadcast channels, and other channels with long distances cause long latency, which naturally prevents error-retry. Thus, channel bandwidth allocation near saturation on real-time imagery streams takes the form of packet collisions. Such packets are most naturally the header/descriptor/payload combination, since each can have its own priority and each forms a constant priority grouping. The constant priority grouping would be the construction used by the originator.

When sharing a finite-bandwidth channel, it may be necessary to pass some data and drop other data. In order for the channel's controlling device to determine fairly which packets to pass and which to drop, priorities for packets might be provided. In many spatial-frequency-based compressed imagery formats such as DCT, subband, and wavelets, the high frequencies represent tiny picture detail that might be lost without much picture degradation. However, the spatial low frequencies, audio, and motion vectors must be heavily protected and may not be dropped without visible artifacts.

The type field would specify the length and format of the priority and/or authorization fields that follow. The length might have 2^type length (2 to the power of the value in the type field, being 2, 4, 8, 16, 32, 64, 128, etc., bytes). Type 0 still represents the absence of the priority field.

Since the priority and their authorization fields will compete at the highest level, it will be necessary for us to define their meaning at the outset. We will further need to define the mappings between the shorter and longer versions of each field.

The format of the fields might be as shown in Fig. A5. The priority field varies from 1 to 8 bytes, allowing very detailed priority levels.

The Bid/Value field allows a packet to have a "bidding price" in a collision with other packets. Such a bidding price would be a value if the price had previously been accepted. A value would imply that tossing the packet would break a contract for delivery to the packet. The meanings of the Bid/Value fields would be tied directly to authorization codes, which would indicate the following:

1. Whether the header was authorized to bid.
2. The "credit rating" (or importance) of the bidder. This could potentially weight the priority field.
3. Whether the bid had been previously accepted, so that the Bid/Value field was the value paid for the payload's delivery. In such a case, tossing the packet would violate the contract. Presumably such a case would occur only when more contracts had been made than were available, so that packets were only tossed by other similar accepted-bid packets with an established value. This is a similar problem to "overbooking" on airlines.
4. Authorization may affect the bid/value. If commissions are paid on some bids or values, and not on others, the net bid or value may differ. This is similar to the problem of bids in different currencies, or direct bookings versus using agents. Thus authorization can indicate the source and/or type of a bid for these purposes.

The priority should be registered in entirety. Thus, the meaning of priority codes might be defined by registration. However, it may be desirable to have priorities take simple linear precedence order, with higher values representing higher priorities. One possible solution is to define the first, or the first and second bytes of the priority to be linear magnitude precedence priority codes. Subsequent bytes, however, could be registered codes, with

unique meanings that are standardized to help resolve priority conflicts.

Other than the potential interactions of authorization on priority and bid/value, authorization can have the following very important uses.

A.7.8 The Authorization Field

Uses of authorization:

1. Pay-per-view target encryption codes (in lock step with receiving system).

2. Channel authorization. For example, is a satellite downlink channel signal authorized for use as a cable head-end?

3. Channel routing authorizations. For instance, are all authorized destinations only on network fork A, such that a source for subnetworks A and B need not carry the payload to subnetwork B. This is the function of supporting a subset of all of the outputs involved in a *Y* connection.

4. End-user authorization for teleconferencing, to indicate who can be included in the teleconference, including who may observe and who may originate.

5. Privacy and protection against unauthorized acceptance or origination of the signal in any use. For example, protection against real-time datastream hackers or unauthorized videophone wire tapping.

6. Authorized user enablement codes. Such codes would authorize user systems for future authorized codes. For example, when a cable subscriber adds a new channel, an enablement code would be sent to the decoder to add authorization interpretation and viewing for the new channel.

7. Diagnostic, statistic, and rating codes for exploring network loads, active users, show ratings, unintentional network disconnects, channel error rates, etc.

8. Copyright information indicating ownership.

9. Copyright fee structures, including where to pay fees.

10. Indications of who may edit a work, whether it may be included in other works, and fee structures for doing so.

11. Possibly an automated mechanism could be constructed to automatically negotiate rights based on prearranged "willing to pay" algorithms, so that clips can be included without undue complication.

A.7.9 Encryption

One function of the transport header is to provide one or more encryption keys for deciphering the payload and descriptor. Many protected users may wish to protect against unauthorized deciphering of the descriptor, since it may contain valuable information that could help in deciphering the payload. Codes could be used for encryption keys, for example, to unlock descriptors. Descriptors, in turn, may contain more elaborate encryption codes to further unlock the payload.

Based upon successful authorization code interactions, encryption codes can be deciphered and applied against the descriptor, the payload, or both. As usual, a type 0 represents that no encryption field is present. Fifteen types are available, with 15 prespecified associated lengths, for encryption. Although the lengths must be prespecified, the meaning of the encryption, or its associated technique, can be completely private. Complex encryption algorithms can be developed and updated between embedded codes in the receiving device, codes in the descriptor, and possibly algorithms transmitted and updated via descriptors.

A.7.10 Sequence Numbers

The sequence numbering field specifies not only packet ordering, but also windows of order and groupings. For example, in some systems various headers and their associated payloads form packets that can update the screen in any order during the frame time before the buffer switch for viewing. However, motion vectors might need to precede compressed image deltas. Thus not only packet grouping, but packet general ordering might be specified.

On some networks, lower-priority packets are delayed, rather than dropped. In such networks, it is necessary for the network controlling mechanisms to understand the bounds on acceptable delay for packets and groups of packets. This field contains a series of codes for defining tolerance of imagery and other real-time streams for being received out of order. Type 0 means no sequence field is present. Fifteen valid type fields with the associated prespecified lengths are available.

A.7.11 Timing Reconstruction

In real-time data streams, it is often necessary to reconstruct precise timing after this timing is disrupted during transport. Timing reconstruction information, concerning the times at which events should occur, are specified in this field. Times can be specified as absolute times, where the transport delays and their bounds are known. Relative times can be specified relative to an arbitrary "start of real-time stream" clock marker, which is set by the receiving device upon receiving the first displayable buffer load.

Synchronization between audio and image, between multiple audio streams, or between streams from multiple sources, is handled via the timing reconstruction field. Resynchronization for removing cumulative jitter effects can also be enabled through the use of this field.

A type of 0 indicates an absence of the timing reconstruction field. The 15 available codes will have prespecified lengths, although their timing meanings may be deferred from some of the types. Of course, each of the types can only receive a single meaning, which meaning must stay in place from then on. The lengths for such unspecified codes must all be specified in advance, however.

A.7.12 Reserved for the Future

This field is unspecified in content and length. Because the total length of the transport header is known, and because this is the last field in the header prior to the pad, this field can maintain flexibility for future use by remaining completely unspecified. All other fields must at least have their lengths specified for each type value.

A.7.13 Pad Bits

Pad bits make the lengths simple for error-correction processing. This is accomplished by making the total of the error-correction/protection field range be the appropriate length for the error-correction format being used. For example, using a 64-bit length type, the length after the error-protection/correction field must be a multiple of 64 bits. Thus if the scope of the protection includes additional transport fields, such as priority and encryption, plus a header and its descriptor, the pad bits would make the sum a proper multiple of 64 bits.

Appendix B — Illustrative Examples of Header Decoding Using "C"

B.0 Background

It is often instructive to represent a design as a computer program written in some appropriate language (in this case C). It verifies the design and provides a basis for comparing the cost and performance of design alternatives. The C language was chosen because it is reasonably universal. Conciseness and consistency are foremost considerations in enabling comprehension and fair comparison; optimal performance is of secondary importance. Optimizations and enhancements would be added in preparation for commercial distribution.

Two programs are described. One decodes a compact header and one decodes an ASN.1 header. They are similar in appearance and use the same basic steps. The primary difference is the compact header decoder selects between multiple formats using table lookups, while an ASN.1 header has only one extensible

format. Each program extracts the block length and standard identifier from the header, and then calls a corresponding function to process the payload.

B.1 Compact Header Decoder

The following program decodes a packet with a compact header. If the packet format is predefined, it calls the corresponding predefined function, otherwise it extracts the standard identifier and block length in a manner similar to the algorithm described in Section 5.2.8. It uses the identifier to look up a decoding function (f), and ignores any blocks with unknown identifiers.

Two table lookups are used to decode the compact header. The length-type table (lt table) contains information used to decode predefined messages and block length. The identifier table (id table) contains information used to decode the standard identifier. One obvious optimization is to combine the two table lookups into a single 256-entry lookup. This reduces the instruction path for some cases, but increases memory requirements.

The identifier is left as a string of bytes used to compute a hash table lookup of a decoding function. If a sovereign state field exists, it is processed together with the standard identifier, but a separate hash table is used. The hash table lookup is performed by the procedure lookup(), which takes identifier address, identifier length, and table selection as arguments, and returns a pointer to the corresponding function (f).

B.1.1 Cautionary Notes

Certain header formats are not yet defined or are reserved for future use. The program below does not support these formats.

Predefined message types have not yet been standardized. To make the code complete a dummy function call, fake(), has been used. When the functions were standardized, the lt table would change accordingly.

Block length is assumed to fit within one 32-bit word. Extending the program and/or the C language to support larger word sizes, thus larger block lengths, is possible and likely to happen as 64-bit processor architectures emerge.

Bit-field ordering and assignment are not yet defined. Choices made in the program below will require further consideration in the context of standardization.

If an unknown identifier is encountered, the lookup function will return a pointer to an appropriate default function that ignores the payload and displays an informative message.

B.1.2 Program Text

The program has two parts — the first part contains table and procedure declarations, the second part (at the end) contains the dozen or so statements actually executed. Throughout the code descriptive notes (comments that are not executed) are placed between comment delimiters (/*...*/).

```
/*Compact header has one of two forms:
*
*Each character in the strings below represents a byte: bytes
*between square brackets are optional: payload bytes are not
*counted
*
*2-byte (minimum) for predefined messages:
*
* "ke[p...p]"
*
*Extended header for longer blocks:
*
* "kel[...l][e...e]i[...i]"
*
*Key:
*
* k :: key byte (length type and id type, presence of a readable
*description)
* e :: error byte
* l :: length byte
* i :: id byte
* p :: payload byte
*
*/
extern void fake();        /*fake predefined func to init lt table*/
extern void null();        /*null func for unused function entries*/

typedef struct {           /*lt table structure*/
   char length;            /*block length or length of length field*/
   void (*function)();     /*predefined function for lt 0..5*/
   char id offset;         /*offset to ident field for lt 6..15*/
}Lt entry;

Lt entry lt table[16] = {                            /*lt table declaration*/
   {2, fake, 0}, {2, fake, 0}, {4, fake, 0}, {5, fake, 0},      /*0..3*/
   {6, fake, 0}, {0, fake, 0}, {1, null, 3}, {1, null, 4},      /*4..7*/
   {2, null, 4}, {2, null, 5}, {4, null, 6}, {4, null, 8},      /*8..11*/
   {6, null, 8}, {6, null, 10}, {0, null, 0}, {0, null, 0},     /*12..15*/
};

typedef struct{            /*id table structure*/
   char length;            /*length of ident field*/
   char table;             /*which table to use in lookup*/
}Id entry;

Id entry id table[8] = {                             /*id table declaration*/
   {1, 0}, {1, 1}, {2, 0}, {3, 0}, {2, 2}, {4, 2}, {8, 2}, {0, 0}
};

extern void (*lookup0)();       /*id lookup function*/


/*EXECUTABLE CODE STARTS HERE*/

int decode compact header (pkt)   /*call standard decoder and
                                    return length*/
   unsigned char *pkt;            /*pointer to packet*/
{
   unsigned char *p = pkt;        /*working pointer to packet*/
   int lt;                        /*length type*/
   int length;                    /*block length*/
   Id entry *pid;                 /*id table pointer*/
   void (*f)();                   /*standard function returned
                                    from lookup*/
   lt = *p >> 4;                  /*get length type from bits
                                    4..7 of key*/
   if(lt < 6)                     /*process predefined formats: lt < 6*/
     {
       length = lt table[lt].     /*get block length from lt table*/
         length;
       (*lt table[lt].function)   /*call predefined function*/
         (p);
       return length;
     }

/*extract enough bytes to cover length field and shift out unused
bits*/

length = ((int)*(p+2) << 24) | ((int)*(p+3) << 16) |
                              ((int)*(p+4) << 8) | *(p+5);
length >>= 32 - (lt table[lt].length*8);

pid = &id table[*p & 0x7];       /*get id table ptr using bits 0..2
                                   of key*/
p += lt table[lt].id offset;     /*move pointer to start of id
                                   field*/
f = lookup(p, pid->length,       /*lookup function*/
pid->table);
```

```
p+ =pid->length                    /*move pointer to start of payload*/
(*f)(p.length-(int)(p-pkt));       /*call function with payload ptr.length*/
return length;
}
```

B.2 ASN.1 Header Decoder

The following program decodes a packet with an ASN.1 header. The ASN.1 header is a compatible subset of the standard ASN.1 EXTERNAL type. As stated in Section 6.4, not all the flexibility of the standard EXTERNAL type is needed to meet the objectives. Thus this decoder has been specialized to support a level of function roughly comparable to that of the compact header decoder.

Decoding an ASN.1 header is performed by parsing a sequence of tokens. Block length is defined by one sequence, standard identification by a second. The header has short or extended forms. The short form is used for blocks of 128 bytes or less and is processed by extracting block length from a single byte. The extended form is processed by constructing block length from multiple bytes.

The identifier is left as a string of bytes used to compute a hash table lookup of a decoding function. The hash table lookup is performed by the procedure lookup(), which takes identifier address and length and returns a pointer to a corresponding function (*f*).

B.2.1 Cautionary Notes

Context-dependent headers are not decoded by the code below. They are decoded by standard decode procedures at a time they are expected. Block length is assumed to fit within one 32-bit word. Extending the program and/or the C language to support larger word sizes, thus larger block lengths, is possible and likely to happen as 64-bit processor architectures emerge. If an unknown identifier is encountered, the lookup function will return a pointer to an appropriate default function that ignores the payload and displays an informative message.

B.2.2 Program Text

The program has two parts — the first contains variable and procedure declarations, the second (at the end) contains the dozen or so statements actually executed. Throughout the code descriptive notes (comments that are not executed) are placed between comment delimiters (/*...*/).

```
/*ASN.1 header has one of three forms:
*
*Each character in the strings below represents a byte; bytes
*between square brackets are optional; payload bytes are not
*counted
*
*2-byte (minimum) header for context dependent messages:
*
* "tl[...l]"
*
*7-byte (minimum) header for short blocks:
*
* "tltli[...i]tl"
*
*Extended header for longer blocks:
*
* "tl[...l]tli[...i]tl[...l]"
*
*Key:
*
* t :: tag byte
* l :: length byte
* i :: id byte
*

/* EXECUTABLE CODE STARTS HERE */

decode_asn1_header(pkt)            /*call standard decoder and return length*/
   unsigned char *pkt;             /*pointer to packet*/
{
extern void ((*lookup())();        /*id lookup function*/
unsigned char *p=pkt;              /*working pointer to packet*/
unsigned int length;               /*block length*/
void (*f)();                       /*standard function returned from lookup*/
unsigned int t;                    /*temp register*/

length=*(p+1);                     /*get 1st length byte*/
if (length<127)                    /*look for short form*/
 {
  t=*(p+3);                        /*get id field length*/
  f=lookup(p+4,t);                 /*lookup function*/
  (*f)(p+t+6,length-
     (p+t+6-pkt));                 /*call function*/
  return length+2;                 /*return total length*/
 }
t=length-128;                      /*calc length of length field*/

/*extract enough bytes to cover length field and shift out excess bits*/

length=((int)*(p+2)<<24) | ((int)*(p+3)<<16) |
                      ((int)*(p+4)<<8) | *(p+5)
length>>=32-(t*8);

p+ =t+3;                           /*move pointer to start of id field*/
f=lookup(p+1,*p);                  /*lookup function*/
p+ =*p+t+3;                        /*move pointer to start of payload*/
(*f)(p.length-(int)(p-pkt));       /*call func (payload ptr and size)*/
return length+t+2;                 /*return total length*/
}
```

Appendix C — Header/Descriptor Task Force Members

Will Stackhouse, Chairman	Jet Propulsion Laboratory
David H. Staelin, Vice-Chairman	Massachusetts Institute of Technolog
Walter Bender	Massachusetts Institute of Technolog
Rita Brennan	Apple Computer, Inc.
Wayne E. Bretl	Zenith Electronics
David C. Carver*	Digital Equipment Corp.
Gary Demos	Demografx
Ephraim Feig	IBM
Branko Gerovac*	Digital Equipment Corp.
Jeffrey W. Johnston	Eastman Kodak Co.
Michael Liebhold	Apple Computer, Inc.
Lee McKnight	Massachusetts Institute of Technolog
Arun Netravali	AT&T Bell Laboratories
Bruce Sidran	Bellcore
D. Scott Silver	Tektronix, Inc.
Richard J. Solomon	Massachusetts Institute of Technolog
David L. Tennenhouse	Massachusetts Institute of Technolog
David L. Trzcinski	PictureTel
Ken C. Yang	Ampex Corp.

*Associate members making significant written contributions.

APPENDIX E

REPORT OF THE TASK FORCE ON DIGITAL IMAGE ARCHITECTURE

Report of the
Task Force on Digital Image Architecture

SMPTE

Society of Motion Picture
and Television Engineers

595 West Hartsdale Avenue
White Plains, New York 10607

Table of Contents

Preface

For some time, the communities participating in the standardization activities of the Society of Motion Picture and Television Engineers have considered the role of television in the future of visual communications. In recent years, the debate has been joined by members of other communities affected by the convergence of communication and imaging technologies enabled by a common set of digital building blocks.

The emergence of digital coding as the common language of visual communications may fundamentally change our view of the world. The extent to which this common language will affect life in the 21st century may be even more profound than the effect that the medium of television has had on life in the 20th century. Television has provided a window to the world – often real-time – for many of the 5.4 billion inhabitants of this planet. This medium of cultural and information exchange has enabled previously isolated populations to join an emerging global village – one increasingly free of barriers. The common digital language offers a unique opportunity to leverage converging technologies, such as television, computers and telecommunications, into a global communications network. Such a network would have the potential to offer a vastly augmented range of services to all system users, thus opening up new markets to all of the affected equipment and service providers.

Worldwide, there is a growing consensus that the time has come to develop standards for television systems based on a new paradigm – appropriate for today – with forethought to future requirements. The introduction of digital technology into imaging industries, together with the widespread introduction of digital communications, creates a window of opportunity to establish a digital image architecture with unprecedented freedom of application and interconnection.

This Report examines some of the fundamental issues that must be addressed in achieving a compatible set of standards enabling a globally interconnected and interoperable visual communications network. The essential concepts for this family of standards include: an open (non-proprietary) system architecture, interoperability, scalability, and extensibility. It is hoped that this Report will stimulate the interest of many groups and organizations involved in the establishment of imaging standards, today and in the future, and lead to agreement on a single system, flexible enough to accommodate a wide variety of needs, while enabling worldwide interoperability.

This Report was prepared by the SMPTE Task Force on Digital Image Architecture and is responsive to the Work Assignment, dated April 1991, which established the following objective:

> To develop and propose a structure for a hierarchy of digital standards to facilitate interoperation of high resolution display systems. The resulting system of standards will be open (non-proprietary), scalable to various performance levels, extensible to new technologies and will be based on the documents submitted by the USA to the October, 1990, meeting of CCIR IWP-11/9, taking account also of the requirements and constraints noted in the joint Advanced Television Systems Committee (ATSC) / Institute of Electrical and Electronic Engineers - Committee on Communications and Information Policy (IEEE-CCIP) meeting on digital system information exchange of March 12-13, 1991.

The Task Force, early in its considerations, identified the need to expand the Work Assignment to include all relevant aspects of digital imaging systems – acquisition, processing, storage, transmission, reconstruction and display – and to consider systems across a much wider range of resolutions than previously planned. This was agreed to and is reflected in this Report. Requirements and constraints noted in SMPTE/IEEE/ATSC cosponsored digital system information exchange meetings have also been incorporated as appropriate.

The Report is, in essence, the outcome of a feasibility study concerning the creation of standards for digital image systems that are scalable and extensible, effecting a high level of interoperability between a diverse range of industries and applications. The work is, as yet, incomplete; however, it has already established an important though preliminary basis for a family of digital imaging standards. The Report raises many new questions and identifies additional work required to refine the concepts that form the basis of a digital image architecture. Of particular importance will be the selection of source and display refresh rates to provide performance and economic compatibility with today's television systems.

The concepts outlined can provide a basis for a modular open system architecture, in which the parameters and characteristics for each module, and the interfaces between these modules, are clearly defined and in the public domain.

Such a system should use common standard components to serve diverse needs across all affected industries. It should enable the movement of image data across application and industry boundaries without degradation and with minimum complication. This is *interoperability*.

Such a system should also provide the ability to adjust image parameters – temporal and spatial resolution, colorimetry and dynamic range – by varying the amount of

data that is stored, transmitted, received, or displayed. This is *scalability.*

A digital image architecture must give forethought to evolution – to incorporate advances in technology within any module, without changes to any other module. It must be backward compatible with today's systems, and forward enabled to accommodate the technology explosions of the 21st century. This is *extensibility.*

This Report was prepared by a Task Force chaired initially by David Trzcinski (PictureTel) and latterly by Dr. Will Stackhouse (Jet Propulsion Laboratory), with a wide participation from the computer, television, post production and telecommunications industries. A detailed list of the membership follows. The Report was considered by the SMPTE Standards Committee at its meeting of August 13, 1992, and subsequently adopted after an in-depth review.

List of Members and Participants

Members

Will Stackhouse (Chair)	Jet Propulsion Laboratory
Walter Bender	MIT
Craig Birkmaier (Editor)	PCUBED
Rita Brennan	Apple Computer
Wayne Brell	Zenith
Barry Bronson (Co-chair)	Hewlett-Packard
Ken Davies (ex officio)	CBC, SMPTE
Gary Demos (Co-chair)	DemoGraFX
Bill Glenn	Florida Atlantic University
Eric Hoffert	Apple Computer
Bob Keeler	AT&T Bell Laboratories
Thomas Leedy	NIST
Peiya Liu	Siemens
Lee McKnight	MIT
Robert Powers	MCI Telecommunications
Tom Meyer	Duir Associates
Alan Reckie	European Community (CCE)
Richard Solomon	MIT
Arpad Toth	Eastman Kodak
David Trzcinski	PictureTel
Michael Wade	DemoGraFX
Ken Yang	Ampex

Participants

Stan Baron	NBC, SMPTE
Si Becker	SMPTE
Rex Buddenberg	Consultant
Robert Burroughs	Panasonic
David Carver	Digital Equipment Corp.
Phil Dodds	IMA
Charles Fenimore	NIST
David Fibush	Tektronix, SMPTE
Paul Fleischer	Bellcore
Branko Gerovac	Digital Equipment Corp.
Barry Gilbert	Mayo Foundation
Christopher Hamlin	Apple Computer
David Herbine	NADC
Clark Johnson	Consultant
Bijoy Khandheria	Mayo Foundation
Edward Krause	General Instrument
Arvid Larson	IEEE-USA
Derrick Lattibeaudiere	Panasonic
Richard Lau	Bellcore
Bernard Lechner	Consultant
Michael Liebhold	Apple Computer
Henry Meadows	Center for Telecommunications Research
Francois Michaud	CBC
Marvin Mitchell	Mayo Clinic
Robert Morrow	USAF Academy
Robert Myers	Hewlett-Packard
Suzanne Neil	MIT
Bruce Penney	Tektronix
Ken Phillips	Citicorp
Ed Post	Quark
Charles Poynton	Sun
Glenn Reitmeier	DSRC
Robert Sanderson	Eastman Kodak
William Schreiber	MIT
Scott Silver	Tektronix
John Sprung	Viacom
David Staelin	MIT
Peter Symes	Grass Valley Group
David Tennenhouse	MIT
Greg Thagard	CST
Mark Urdahl	IBM
John Weaver	Liberty Television
Merrill Weiss	Consultant

International Participants

Norbert Gerfelder	Fraunhofer Computer Graphics, ISO/IEC
Rainer Hofmann	Fraunhofer Computer Graphics, ISO/IEC
Detlef Kroemker	Fraunhofer Computer Graphics, ISO/IEC

1 Executive Summary

The SMPTE Task Force on Digital Image Architecture was charged with developing and proposing a structure for a hierarchy of digital image standards that would facilitate interoperation of image systems. The major objective was to establish the basis for image systems that are open, scalable and extensible, thus meeting the perceived needs for image communications in the environment likely to exist as computers, television and communications converge, enabled by pervasive digital technology.

The Task Force, formed from representatives of the affected industries and applications, has examined the issues, setting out those that are believed critical at this time, and has modelled, for discussion, further refinement and testing, one possible approach that meets the basic requirements. It has also produced extensive tutorial information concerning the matters under consideration.

The key concepts of the approach are defined in Section 2, setting the conditions for image systems that are:

- Open – the modules and interfaces forming the architecture are fully defined and in the public domain
- Interoperable – images and related equipment may move freely across application and industry boundaries

Such systems would be based on a hierarchy that is:

- Scalable – supports a wide range of image capabilities
- Extensible – future-proof to the extent possible
- Compatible – supports existing television practices and standards when possible

Current and future image systems are presented and analyzed in Section 3 of the Report, which also states the main objectives of the Task Force activity:

- To establish the fundamental properties of image systems
- To examine the technological trends with a view to a prediction of future capabilities
- To arrive at architectural guidelines that will achieve the objectives of interoperability, scalability and extensibility

Section 3 of the Report also establishes the fundamental concepts upon which a model for an open digital image architecture can be constructed, taking into consideration the objectives defined above.

Section 4 details the critical issues in the development of a suitable image architecture meeting the stated objectives:

- The establishment of scalable and interoperable hierarchies for basic image parameters
- The selection of a family of image acquisition rates and related display refresh rates based on a progression that permits display refresh at integer multiples of the acquisition rate. Backward compatibility, if required, to the image acquisition rates currently in use (24, 50, 59.94 and 60) should be accommodated in the design of the standard modules which will interface these existing systems with the digital image architecture.
- Use of a square sampling grid as a simple and effective common expression of images
- Selection of analysis and coding schemes for color and luminance that would allow useful and effective scaling of this image data while maintaining a high level of interoperability
- Coherent sampling of the image based on the use of progressive scanning techniques
- Use of headers/descriptors to identify the content and conform to the characteristics of the data stream
- The establishment of appropriate levels of compatibility with current television and motion picture standards

It is believed that this approach will result in systems that achieve a good level of compatibility with current television and imaging systems, while placing a minimum of constraints on the path to the future (extensibility).

A model of an open architecture approach to image standards is developed in Section 5, one that is both compatible with the present and extensible to the future. It is based on a low order hierarchical approach, using image tiles. The model defines four levels of resolution and takes account of a number of possible aspect ratios currently in use. Additional analysis is provided regarding the selection of an appropriate family of image acquisition rates and display refresh rates. Finally, a scalable coding approach is proposed that offers the ability to produce image data in packages that can be combined to produce images at a variety of spatial and temporal resolutions.

The Task Force Report is expected to be of interest across a wide range of industries and applications. Section 6 examines the industries likely to be most affected, their specific imaging needs and the possible impacts of a defined digital image architecture.

In Section 7 the Task Force suggests additional work that must be completed to move towards a full implementation of a digital image architecture. The list of suggestions included in Section 7 is not exhaustive; it is recognized that in the process of validating the architectural concepts, additional areas for further analysis will be identified. An extensive list of questions is included which should be considered in the process of establishing standards for an architecture.

The suggestions include the following items of high priority:

- Development of strawman applications in a few critical industries with a view to validating the concepts of an architecture
- Detailed analysis of potential breakthroughs in image technologies to establish a solid basis for extensibility
- Analysis, simulation and optimization of proposals concerning the selection of a family of image acquisition rates and related display refresh rates
- Detailed analysis of the possible routes to colorimetry and dynamic range expressions that allow scalability and interoperability. This work will require simulation and subjective evaluations to validate the results.

A considerable amount of background and tutorial material was developed during the preparation of the Report. Some of it is believed to be of value generally or for reference in future work on the development of the digital image architecture. This material is included in Section 8:

- A glossary of terms
- An analysis of some of the issues surrounding temporal rate conversions

2 Key Concepts

2.1 Introduction

As a starting point in the process of developing and communicating the requirements for a digital image architecture, it is important to establish a clear definition of the key concepts upon which the architecture is to be based. In many cases, existing definitions must be enhanced to bridge the gap between current practice and future requirements embodied in the architecture.

Two reference documents were utilized in the process of creating the definitions which follow:

- The IEEE Dictionary, ANSI/IEEE Std 100-1988
- The Final Report of the SMPTE Task Force on Headers/Descriptors, January 3, 1992

Definitions obtained from the IEEE Dictionary are presented in "quotations" – they provide a reference point for the expanded definitions developed by the Task Force. Definitions presented in the Report of the SMPTE Task Force on Headers/Descriptors proved to be incomplete for the needs of this Report, due to the expanded Work Assignment for the Task Force on Digital Image Architecture. While the definitions in this Report are consistent with the earlier work of the Task Force on Headers/Descriptors, they provide an expanded understanding of the key concepts for a digital image architecture.

2.2 Digital Image Architecture

A system architecture defines "the structure and relationship among the components of a system."

One of the major objectives of this Report is to define a system architecture which promotes sharing of images and equipment across applications and industry boundaries. To achieve this goal, the digital image architecture must be highly flexible to deal with a variety of diverse requirements, including the evolution of technology.

A digital image architecture should be an open system, that is, one made up of functional modules with standard, public interfaces which can be assembled into a functional system – "a set of interconnected elements constituted to achieve a given objective by performing specified functions." Explicit objectives of the architecture include:

- Backward compatibility with existing systems
- Ability to incorporate technological improvement and evolution over time
- Risk reduction (mitigating the impact of incorrect predictions). Individual issues, such as those discussed in this Report, may be encapsulated into discrete building blocks which can be changed without affecting the rest of the system.

The digital image architecture supports both natural and synthetic imagery including:

- Text
- Graphics
- Animation
- Still images
- Motion image sequences
- Audio
- Related image data

A key feature of the architecture is that it allows decoupling of the system into functional modules. The functional modules of the architecture are:

- Acquisition or generation – including processing (transformation, enhancement, composition, interpretation) and storage
- Transmission – including compression and storage
- Display – reconstructing the image for the human visual system, including hard copy print

2.3 Interoperability

Interoperability is the sharing of images and equipment across application and industry boundaries. When dealing with digital image representations, this sharing should be facilitated without degrading image quality due to transformations in temporal and spatial resolution, grid geometry, and image aspect ratio.

This requires careful attention to the definition of the interfaces – the shared boundaries – between the functional modules.

The key interface definitions are:

- A data definition between the acquisition and display modules
- A service definition between acquisition and display modules to the transmission module

2.4 Hierarchy

A hierarchical digital image architecture is one in which various levels of performance are supported:

- A variety of spatial resolutions, aspect ratios, and temporal rates

- A variety of colorimetry and dynamic range specifications
- A variety of image encoding standards.

The architecture is hierarchical in order to address the requirements for scalability and extensibility.

2.4.1 Scalability

To scale is: "To change the quantity by a factor in order to bring its range within prescribed limits."

Scalability deals with the ability of an imaging system to adjust the level of performance by varying the amount of data that is stored, transmitted, received, or displayed – up to the maximum resolution that was originally acquired. A number of specific definitions are implied:

- Transmit Scalability – the ability to encode a visual sequence at various spatial and/or temporal resolutions to conform to specific transmission requirements. These requirements may include support for both fixed and variable bandwidth channels, and specific transmission standards. Transmit scalability includes scalable compression – the ability to encode a visual sequence so as to enable the decoding of the digital data stream at various spatial and/or temporal resolutions. Scalable compression techniques typically filter the image into separate bands of spatial and/or temporal data. Appropriate data reduction techniques are then applied to each band to match the response characteristics of the human visual system, the capacity of the transmission subsystem and the economics of the display subsystem.
- Storage Scalability – the ability to employ storage devices of different capacities and access speeds
- Receiver Scalability – the ability of a decoder to extract, from a single data stream, only that data required for displaying a portion of the image and/or a reduced resolution image, either spatially or temporally
- Display Scalability – the ability of a display to conform to visual sequences of varying spatial and/or temporal resolution. With scanning display systems this may be accomplished by changing the synchronization rates for the scanning process, or by creating a window that utilizes an appropriate subset of the available pixels.

2.4.2 Extensibility

Extensibility in the design of a hierarchical digital image architecture allows the system to evolve with advances in the underlying technologies so that additional levels of performance can be implemented, without rendering obsolete those existing products that conform to the basic requirements of the imaging hierarchy.

Extensibility implies designing evolution into the system. The transmission and display modules of the system should be cast as building blocks. The building blocks, because of their inherent modularity, may freely evolve over time.

3 Analysis of Imaging Architectures

3.1 Establishing a Framework for Analysis

For more than two decades, the application of digital processing techniques has contributed to the evolution of analog composite television systems, especially in the areas of video recording, image processing, and image synthesis. This evolutionary use of digital technology had little effect on the perception of imaging systems: from this perspective many observers believed that digital video would gradually replace analog video without any fundamental changes to the foundation of imaging systems.

However, in the past few years this evolutionary view of imaging systems has been challenged. At the 26th Annual SMPTE Advanced Television and Electronic Imaging Conference, John Watkinson suggested that we analyze the impact of digital technologies from another perspective: "To think that digital technology only impacts the underlying equipment and that otherwise it's business as usual is to miss the larger transformation that is occurring in each of the affected industries."

From Watkinson's perspective, the transition to a new digital imaging architecture represents the opportunity for a new paradigm. Proponents of this position have encouraged system designers to step back and take a global view of the impact that digital technologies are having on every industry that deals with electronic imaging: to think not just in terms of delivering ever-improving levels of image quality, but of what being digital really means.

John Naisbett in his 1982 best seller *Megatrends: Ten Directions for Transforming Our Lives*, stated that new technologies go through three phases as they become part of our daily lives. Applying Naisbett's model to the evolution of electronic imaging systems leads to the following three paradigms:

- Paradigm 1 – the technology is threatening, so it enters society in a non-threatening way: i.e., entertainment in the form of broadcast television
- Paradigm 2 – the technology is extended to improve products that we already use or to simplify the work that we do: i.e., professional video, camcorders and the home VCR
- Paradigm 3 – the technology enables entirely new applications: i.e., the integration of video with other media and participation by the consumer

From the new perspective, being digital deals with the shift to the third paradigm. It is the enabling technology that has made it possible for this Task Force to analyze the requirements for interoperability, scalability and extensibility, and to propose a set of guidelines to accomplish these goals. What are the aspects of being digital that have brought about this transformation in perspectives?

A major factor has been the geometric progression in computer processing capabilities – doubling computational power every two years, with little change in cost or size. This progression is projected to continue well into the next century. As a result, high-resolution still-image processing capabilities are now within reach of every computer user. Techniques once reserved for high-end workstations are now commonly applied in desktop computing, including the recent addition of full motion video as a data type.

Video has also been a major beneficiary of the technology progression. Production systems that only a decade ago required a six foot rack of electronics can now be implemented in a few rack units – or on a few cards that plug into a personal computer.

The tremendous increase in computational power has enabled another critical aspect of being digital – video encoding based on the use of digital compression techniques to reduce the required data rate. A variety of compression technologies have evolved that remove image redundancy within and between video frames. The required data rate may also be significantly reduced by more efficient coding of the image at the source. Developments of such techniques are progressing rapidly and may become useful in the near future.

While compression technology has existed for many years, and continues to evolve, practical implementations for video have only become possible in the past few years due to the rapid evolution of digital processing technologies. This in turn has simulated new research into scalable video encoding techniques that will allow multiple levels of image quality to be extracted from a single image data stream. Some observers predict that the processing power required for the decoding of scalable digital video streams will be universal and inexpensive before the end of this decade.

Improvements in data compression perform the same function as increases in bit carrying capacity in the communications system – delivery of more bits to the user. In the past decade, increases in communications capacity of several orders of magnitude have occurred.

In such an environment, the longevity of new equipment purchases may be dependent upon a digital image architecture that is designed with adequate provisions for extensibility. To meet this objective the Task Force has focused its attention on three areas:

- Fundamental properties of human visual perception and their effect on the design of imaging systems (Section 3.2)
- Technology trends that can be predicted with reasonable certainty (Section 3.4)
- And (most important) architectural guidelines that if held constant will promote interoperability, scalability and extensibility (Section 4)

3.2 Properties of Human Visual Perception

The human visual system deals with the physical world both in terms of its ability to resolve image detail (spatial resolution), and changes in the environment (temporal resolution). We experience the world visually by capturing light directly from a source, or as the reflections of light off of objects in our physical environment. The resulting perceptions of the environment are typically described in terms of size, shape, brightness, color, depth, direction, and speed. These qualities arise in the brain's image processing circuitry; essentially they result from a comparison of the acquired visual cues with what we have learned about the world's intrinsic structure.

As research has revealed more about the physiology of vision, prevailing theory has evolved, placing major emphasis on the computational and cognitive role played by the brain and local image receptors. In turn, this research is providing potentially valuable input to the designers of digital imaging systems.

3.2.1 Human Image Acquisition

The human visual system relies on multiple image receptors to deal with the diversity of environments that it encounters: cones are utilized for color image acquisition over a wide range of illumination levels; rods are utilized for monochrome image acquisition over the lower range of illumination levels.

The eye contains approximately two million cones and 120 million rods. The cones are organized into three broad groups of receptors that are sensitive to light in specific spectral bands; while these bands have significant overlaps, they roughly conform to the red, green, and blue portions of the spectrum. Red and green receptors each outnumber blue receptors by a factor of two to one. The dispersion of these receptors is not uniform, thus spatial perception deals with a complex matrix of receptor types and cognitive processing by the brain.

The center of the visual field, an area called the fovea, contains 30,000 to 40,000 cones and no rods. Outside the fovea the density of cones diminishes, interspersed among the high density rods. The cones within the fovea are responsible for high spatial detail perception while the extrafoveal cones and rods play an important role in visual search and influence directed eye movement. Central vision enables us to see detail, while peripheral vision is attuned to change.

Although high spatial resolution vision is restricted to the fovea, the visual system acquires high resolution images over a wide portion of the field of view. This is achieved through involuntary eye movements: high frequency tremor, slow drift, and rapid saccade.

Research has determined that it takes several hundred milliseconds for the eye to acquire a high spatial resolution image, synthesized from a number of overlapping views. Slow drift and rapid saccade are the mechanisms used for repositioning the fovea to acquire these multiple impressions. The tremor appears to be a mechanism to remove high frequency spatial noise. The tremor's oscillation occurs at a frequency range of 40 to 80 Hz over an area approximately equal to the size of a single cone.

Since human vision is binocular, involuntary eye movements also contribute to depth perception; the brain processes these overlapping views to obtain differences from which depth and spatial properties are inferred.

The spatial resolution of moving objects is also linked to eye movement:

- If the eye does not track a moving object, the object will move across multiple sensors creating an image with low spatial resolution. In this case resolution depends on the rate of movement; rapid motion will create a blurred response.
- If the eye tracks the moving object, receptor stimulation is aligned with the object and additional samples are collected over time. Thus the image will be of higher spatial resolution.

3.2.2 Human Visual Processing

Much of the research in visual science today is focused on the processing of data acquired by the image receptors. A variety of specialized analyzers in the eye process data from small localized regions and accumulate the results into channels which are processed by the brain to create an integrated view of the physical environment.

There is evidence that the brain directs the activity of the image receptors for processes such as establishing white balance and light sensitivity levels. Simple localized analyzers are used to enhance the data transmitted back to the brain. Some of these analyzers are sensitive to a particular edge orientation; there are sufficient analyzers at each location to represent a full set of edge orientations. Additional tuned analyzers cover portions of the range of human sensitivity for spatial frequency, spatial position, temporal frequency, direction of motion, and binocular disparity.

The data processed by these analyzers moves to the brain through two types of channels: a set of fast responding

channels with relatively transient responses to stimuli, and a set of slower channels with relatively sustained responses to stimuli. Transient channels process the output of analyzers that are tuned for low spatial and high temporal frequency stimuli. Sustained channels process the output of analyzers that are tuned for high spatial and low temporal frequency stimuli.

3.2.3 Thresholds for the Perception of Flicker

Above certain frequencies, flickering light sources will appear as a continuous light source. The relevant frequency is called the critical fusion frequency and varies with the level of illumination. Separate flicker thresholds exist for the transient and sustained processing channels.

Transient channels are sensitive to flickering light sources with low spatial resolution; this type of stimulation appears as wide-area flicker and is most noticeable in peripheral vision. At low levels of illumination (where rod vision is used) flicker fusion occurs at frequencies of only a few Hz; as the level of illumination increases and cone vision is triggered the fusion frequency increases.

Flicker from low light level sources such as a television or movie screen typically disappears in the range of 20 to 60 Hz. As screen size increases, taking up a larger portion of the field of vision, or if screen brightness increases, the frequency for flicker fusion increases.

Sustained channels are sensitive to flickering light sources with high spatial resolution; this type of stimulation appears as small area-flicker, often associated with moving objects. In this case the flicker fusion frequency can be much higher than for wide-area flicker; this form of flicker manifests itself as strobing of the object.

An excellent example is found in the single pixel horizontal lines often used in computer graphics. These lines do not appear to flicker on a progressive scan computer display which is refreshed at rates above 60 Hz; but if the same image is presented on an interlaced video display the single pixel lines are presented in every other field (at 30 Hz) and they flicker. This is due to the fact that the persistence of the display phosphor is of shorter duration than the refresh rate; higher scanning rates (either progressive or interlaced) eliminate the flicker.

3.2.4 Tuning Electronic Imaging Systems to Match Human Visual Perception

Our improved understanding of human visual perception together with an exponential improvement in electronic image processing techniques has set the stage for the design of a new digital image architecture.

In order for a new digital image architecture to be interoperable it must deal with existing imaging technologies. This requirement can place many constraints on the design of the architecture. It is important to understand the reasons that these constraints exist to determine if the new architecture must be similarly constrained.

3.2.5 The Elimination of Flicker on Scanning Displays

Information is presented in Section 4.3 which suggests that the refresh rate for scanning CRT displays should be linked to the field of view and brightness of the display. Lower refresh rates are acceptable when the display covers a narrow field of view, as is the case with our existing analog composite video delivery systems. Lower refresh rates are also acceptable for a display with a wide field of view at low brightness levels; typically this type of display requires a viewing environment with low ambient light levels such as a theater.

As the display covers a wider field of view at higher levels of brightness, the refresh rate must be increased to eliminate wide-area flicker. If information with high frequency edges, such as computer generated text and graphics, is presented on the display it must also be refreshed at a higher rate. The computer industry uses progressive scanning with refresh frequencies above 60 Hz to eliminate flicker; larger displays (≥ 16 inches diagonal) are typically refreshed at 72 or 75 Hz.

The same requirements for the elimination of wide-area flicker are now starting to influence the development of display systems for home entertainment. At the higher end of the home entertainment market it would be desirable for displays to provide a 50 degree field of view, and be viewable at normal room ambient light levels. Such a display has resolution and refresh requirements nearly identical to a large personal computer display.

3.2.6 Constraints that Dictated the Use of Interlace Scanning Technologies in Analog Composite Video Systems

Several factors influenced the decision to use interlace scanning techniques for acquisition and display when our composite video systems were designed:

- Bandwidth available for signal processing and transmission
- The elimination of wide-area flicker

Both interlaced and progressive scanning were evaluated; interlace proved to be the best solution to reduce signal bandwidth and minimize flicker in the display.

3.3 Models for the Design of an Imaging Architecture

As the design of a new digital imaging architecture is approached, it is important to take into account of all the

applications and industries that may utilize the architecture, as well as the economic contributions of each in the development and purchase of the system components (see Section 6). Experience with analog television has amply demonstrated the value of "economies of scale." The opportunity now exists to design an open digital imaging architecture that is based on generic, inexpensive, and increasingly powerful processing elements.

The choice of a digital image architecture has implications that reach far beyond the normal realm of standards-setting activities. Telecommunications, television, and computing have made major impacts on life in the 20th century – their integration is likely to have a profound affect on the way that the world communicates, is educated, works, plays and relaxes in the next century.

3.3.1 Components of the Model – Resolution

The level of resolution perceived by a viewer is a function of the distance of the viewer from the display. Thus, to design a digital image architecture that provides constant perceived resolution across applications that involve different viewing distances (e.g., close for a computer display, further away for a conventional TV screen, still further for a large flat-panel display), the system must be scalable in terms of image resolution.

In addition to holding perceived resolution constant under varying viewing distances, it is considered desirable to provide even greater resolution in some applications, as discussed below and as implemented in current proposals for advanced television systems.

While it would be desirable to design an imaging architecture in which resolution could be scaled in a continuous fashion, a hierarchy based on a progression of related image resolution levels can provide similar benefits to system designers and simplify the process of interoperation. Sections 5.2 and 5.3 provide a detailed analysis of the variables that affect the perceived resolution of a display and illustrate the principles of a hierarchical digital image architecture with a progression of four image resolution levels.

Throughout this Report, the concept of a multi-resolution hierarchy will be discussed and refined. The Task Force has constructed a model to facilitate this discussion. It is recognized that many different sets of numbers can be used within this model. Four levels of resolution have been identified and defined; additional levels can be added to the progression, as enabling technologies allow support for higher levels of resolution. The four levels in the model are:

- Low Resolution – systems designed to deliver images over a close- up, narrow field of view, particularly when constrained by cost or bit rate, may deliver imagery at less than average human visual acuity. Videoconferencing, videomail and personal entertainment systems are good examples of applications for low resolution imaging.

- Normal Resolution – systems designed to deliver images with resolution requirements that approximate average human visual acuity. This level of resolution is suitable for the delivery of natural imagery, as is the case in most entertainment applications. Present television systems are in this range, and studio quality analog component video is at the high end of the range.

- High Resolution – systems designed to deliver images over a wide field of view, with high spatial and/or high temporal frequencies, typically have resolution requirements near the limits of human visual perception. Applications include motion picture delivery in theatres, entertainment and information presentations to large audiences, and personal computer displays.

- Ultra High Resolution – systems designed to deliver images over the entire field of view (or beyond), at the spatial and temporal limits of human visual perception, place the most demanding requirements on an imaging architecture. Theatre delivery systems such as Showscan® and IMAX® are in this category. Electronic imaging systems for simulation are reaching this level of resolution as the enabling technologies make delivery at this level feasible.

The concept of interoperability first appeared in the early days of television because of the need to integrate film material into television program content. Unfortunately, film and video were in some respects incompatible. Elaborate shuttering mechanisms were developed for the telecine to make it possible to display film in the world of video; thus the concept of interoperability was born. For 525/60 the compromise was the use of 3-2 pull down, to accommodate the change from 24 to 30 fps (frames per second). The solution for 625/50 was easier – a 4% speed change, playing programs acquired at 24 fps at the television rate of 25 fps.

The evolution of electronic image acquisition systems has been driven primarily by the mass market transmission standards – NTSC, PAL and SECAM. New applications for video such as professional and personal video systems have been enabled through the economies of scale associated with these standards.

Thus, applications which required higher resolutions than those offered by NTSC, PAL and SECAM have either been forced to bear the expense of system development and low volume manufacturing – a luxury primarily reserved for the military – or to wait for the next imaging standard to evolve. It is interesting to note that the equipment developed for the various analog HDTV systems has seen extensive use in professional applications that need the added resolution afforded by these systems.

3.3.2 Components of the Model – Acquisition, Transmission and Display

The path through which an image passes from capture to display may involve as many as five major steps as shown in Figure 3.1. These steps are discussed in detail in Section 6.2.3 of this Report. As the imagery moves from one step to the next, it may be stored at one of several quality levels.

The first two steps are

1. Capture (by an image sensor)
2. Processing

and typically require production quality storage to preserve as much of the original imagery as possible for subsequent processing. After the processing steps have been completed, the imagery may be stored at a lower level of quality for release to the distributor of the imagery; this is often referred to as contribution quality storage.

Delivering the imagery to the consumer typically involves the third step.

3. Transport

The imagery may be encoded and stored at a lower level of quality to conform to the transport characteristics; we refer to this as distribution quality storage.

Finally, the imagery must be decoded for display, requiring

4. Reconstruction
5. Display

The consumer may store the imagery for viewing at another time; this also requires distribution quality storage.

Some of these steps tend to be grouped with a specific level of storage quality, as illustrated in Figure 3.1. This allows a further simplification of the model based on three major system components – ACQUISITION, TRANSMISSION, and DISPLAY.

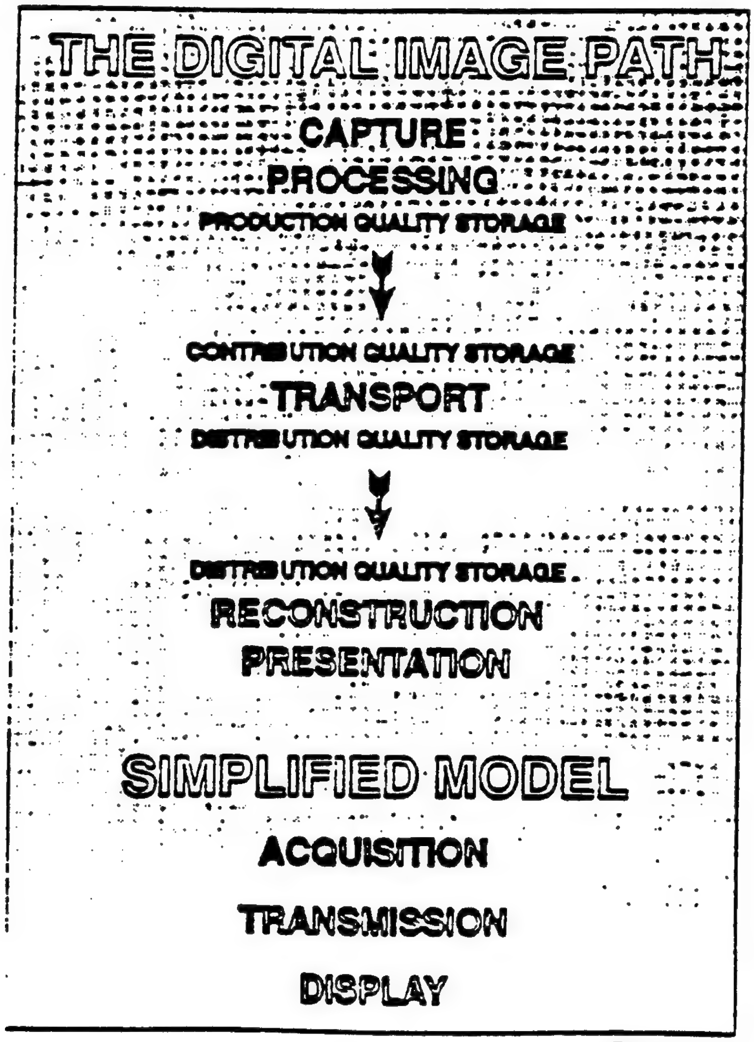


Figure 3.1

3.3.3 A Closed Architecture Model – Analog Composite Video

The transmission standards for the existing composite video systems frequently require all of these components to operate in close synchronism. The display is synchronized with live or taped program material that feeds the transmission system. Imagery acquired at other spatial or temporal resolutions requires conversion into the spatial and temporal specifications of the transmission standard. Such an architecture is depicted in Figure 3.2.

The advent of video recording provided a degree of decoupling of acquisition from the other components, allowing program producers to create program content without real-time constraints; however, transmission and display remain tightly coupled. Recording media for program content have typically been coupled to the transmission standard to take advantage of the bandwidth reduction techniques applied in the system. The design of consumer VCRs is based on compatibility with the transmission standard; packaged media played by the VCR must therefore conform to the same standard.

While interoperability between the various analog composite video systems has had to overcome differences in frame and line rates, these systems have been remarkably extensible. The acquisition, transmission and display components and the associated services of the system have evolved continuously over the past fifty years.

With the introduction of analog component video recording and processing systems in the '80s the video industry took a major step toward completely decoupling acquisition from transmission and display. The production community soon discovered the advantages of this decoupling.

By using analog component equipment for both acquisition and production, it became possible to edit video without concern for the multi-field color framing sequences that exist in subcarrier encoded composite video systems. Producers also discovered that fewer artifacts were introduced when layering video using component vision mixers and digital video effect systems. Decoupling of acquisition and production equipment from the encoded transmission standard produced far better results than could be achieved with composite video acquisition and production equipment – and the same video recorders also produced encoded outputs for transmission of the program.

3.3.4 An Open Architecture Model – Digital Hierarchies

In the '80s, the publishing industry experienced the collision of analog and digital technologies. Today, inter-

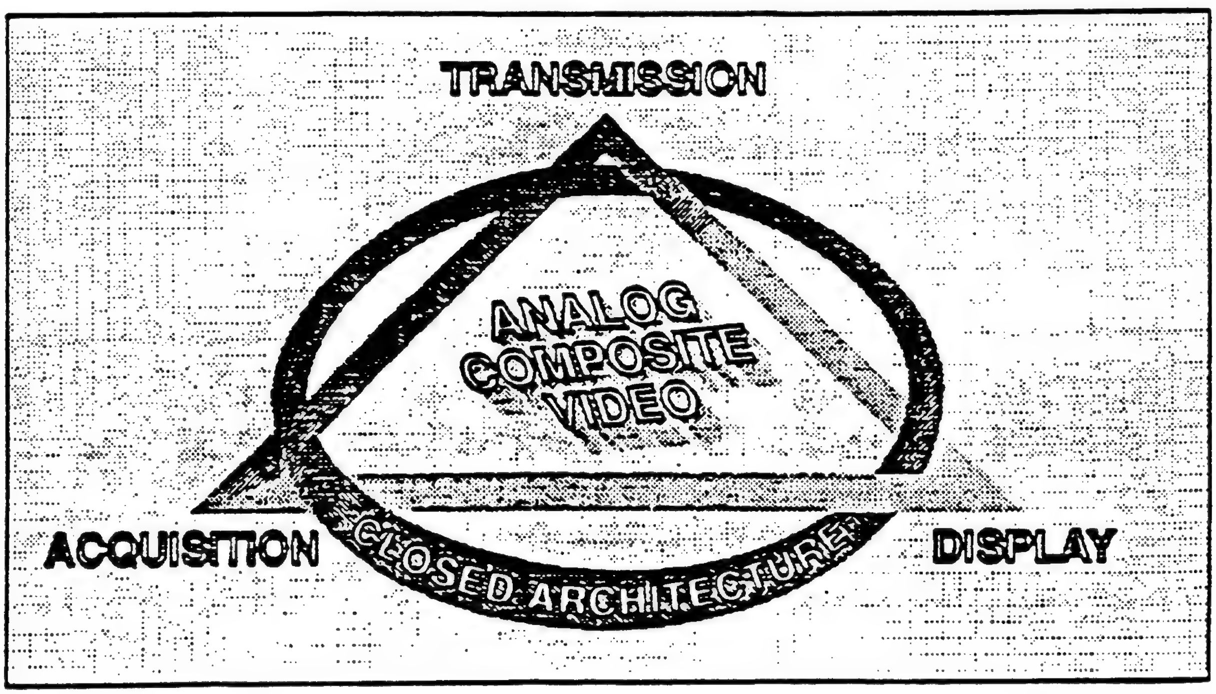


Figure 3.2

operability of media in the publishing industry is the rule rather than the exception, as digital image and document processing techniques, generally categorized under the umbrella of Desktop Publishing, have replaced traditional analog techniques.

To a large extent, the transition from the analog representations of printed media – type, line art, halftones, and color separations – to their digital counterparts, has been enabled by the use of scalable hierarchies for the acquisition, transmission, and display of printed materials. The tools for acquisition and production of print media have been separated from the display hierarchy, allowing output at the desired level of resolution.

Electronic transmission is also beginning to play a major role in the publishing of documents. Compact representations of printed media, using page description languages, have allowed high quality print representations to be moved efficiently through the telecommunications network using low data rate modems. Remote printing of documents on fax machines or networked printers is commonplace.

The desktop publishing metaphor has been used as a model to predict similar transitions in other media industries, most notably Desktop Video. However, the transition has not occurred at the pace that many industry pundits have predicted. This is due, in large part, to the difficult task of breaking the problem up into manageable components: to create separate hierarchies for acquisition, transmission and display of motion imagery.

Interoperability of video systems with other media is facilitated by a complete decoupling of the acquisition, transmission and display into separate hierarchies for each component. Such an architecture is depicted in Figure 3.3. Scalable representations of video will be enabled by this decoupling, and technological advances in one hierarchy can take place without upsetting the apple cart in the other two.

If a hierarchical digital imaging architecture is used as the model, a digital advanced television system can be implemented that is equally adept in delivering low cost solutions that conform to a single hierarchy, as well as more expensive scalable solutions that support multiple hierarchies.

The acquisition hierarchy can provide image capture solutions at various price/performance points that are appropriate for the application. Production systems can evolve that deal with single image formats, or multiple formats within the hierarchy. This is of particular importance to producers of program content with significant archival value. Imagery can be captured at a higher level

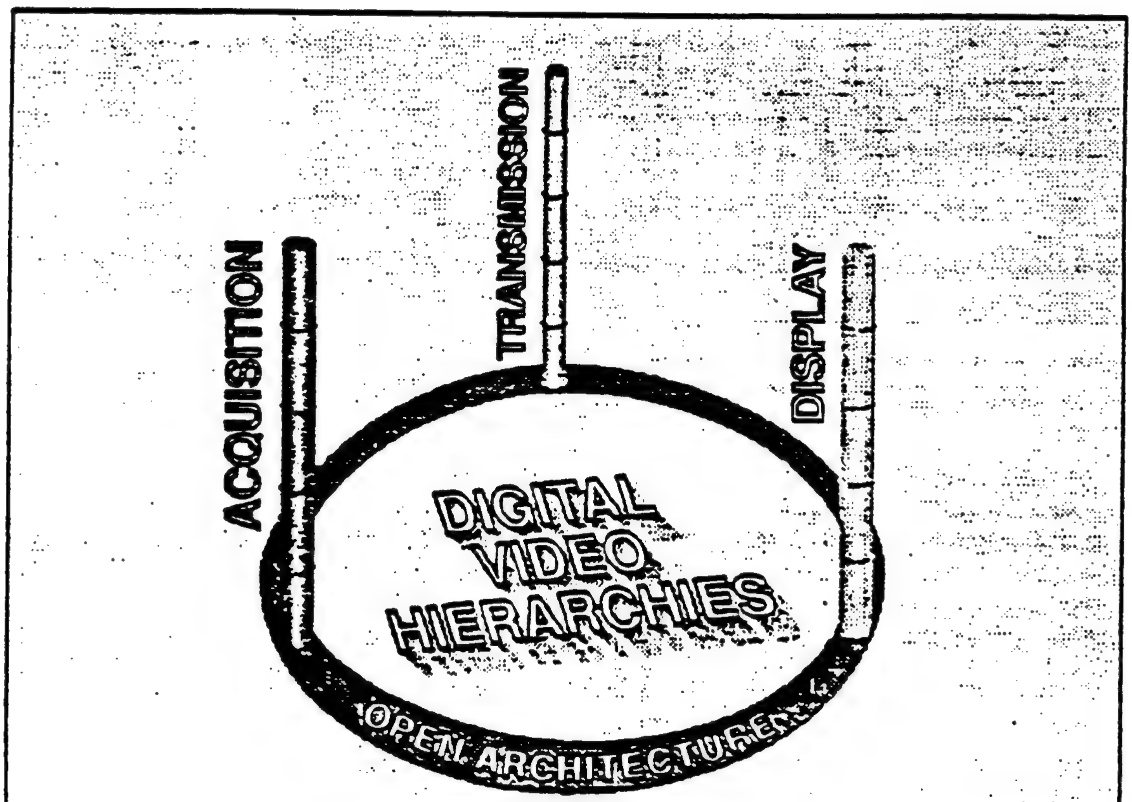


Figure 3.3

in the acquisition hierarchy with an eye toward distribution at one or more of the lower levels of the transmission hierarchy; the archival value of the program is protected since it can be released at higher quality levels in the future as consumers purchase products at a higher level in the display hierarchy.

Viewing transmission as a hierarchy is critical to the concept of interoperability. A hierarchical imaging architecture would support a progression of image quality levels that are interoperable and extensible, and allow for incremental improvements in image quality within a single transmission standard. This requires the use of a scalable encoding structure: a core image would be encoded at the first level of the hierarchy, and enhancement information would be encoded for each of the higher resolution levels supported by the transmission standard.

A scalable encoding structure may be more difficult to design and possibly less efficient for a given quality level than an encoding designed specifically for that level. It has, however, several advantages that will accrue over time:

- It allows receivers to be designed at different quality (and cost) levels and still share components and a common transmission
- It could alleviate the need to create a new transmission standard from scratch as the evolution of compression algorithms permits improvements in quality (for the same bit expenditure)
- It can conserve bandwidth when image quality is added incrementally to existing transmissions, rather than creating new transmissions

The economic benefits associated with scalable image encoding will be significant. The emerging consensus among experts in video compression technology is that scalability will carry a minor penalty for encoding overhead. Consider the impact on media server storage systems: a single scalable representation will make more efficient use of storage than multiple copies of the same material at different scales.

The display hierarchy allows for a variety of products to evolve at various price/performance points that are appropriate for the application. Some display systems will evolve to single performance levels while others will offer multiple levels of performance within the transmission and display hierarchies.

Scalability plays a major role in the design of decoder and display components. If the transmission system delivers a scalable payload, only that portion of the information which is required for the display system need be decoded. A small personal information system may only need the low resolution component while a high-end home entertainment system can utilize all of the resolution components.

3.4 Factors That Have the Potential to Fundamentally Change Digital Image Architectures

Real world constraints, especially with respect to cost versus performance, are the driving factors in the implementation of a digital image architecture. In determining the requirements for the architecture, the Task Force has analyzed the current market situation as well as technology and regulatory trends that may reshape the market in the next few decades.

- The technology for image acquisition is still evolving rapidly – cameras capable of delivering the highest levels of spatial and temporal resolution are expensive. It is anticipated that the image acquisition market will continue to be characterized by a wide range of products at different price/performance levels.
- The storage of image data for subsequent processing or viewing is a major component of the expense in most imaging systems – even with digital video compression. Storage cost is directly related to the data content of the image – higher performance equates to higher system cost.
- Transmission channels such as terrestrial and satellite broadcast are limited by regulation and practical technology. Networked systems (cable and telco) are limited by the installed infrastructure and the cost of providing additional bandwidth for new applications; while available bandwidth has increased by many orders of magnitude, many additional orders of magnitude must be achieved to make the delivery of high resolution imagery as generic and affordable as our telephone service is today.
- When viewing imagery, a number of inter-related variables affect both the cost and performance of the display system: viewing environment (especially ambient light), viewing distance, field of view, and the desired resolution. It is anticipated that the display market will continue to be characterized by a wide range of products at different price/performance levels.

3.4.1 A Shift in the Pricing Structure of Broadband Telecommunications

Changes in the regulatory climate are likely to cause increased competition among all networked service providers (telcos, cable, data networks, etc.) and encourage service providers to upgrade the quality and capacity of these networks.

The current pricing structure for broad band telecommunications is typically based on channel bandwidth – the purchaser uses and pays for the entire channel regardless of the amount of information moved through it. In the

future, greatly increased channel bandwidth and packetized encoding schemes using headers/descriptors for packet identification will cause a shift in pricing structure – the purchaser will pay only for the information content that moves through the channel. This concept when applied to video services has been described as pay-per-view-per-bit.

This shift in pricing structure is likely to act as a catalyst for the rapid evolution of video compression techniques and transmission standards, with an emphasis on two areas:

- Data rate reduction – delivering increasingly higher levels of quality while holding bit rates constant
- Scalability – a shift toward encoding schemes that allow for different levels of image quality to be extracted from the encoded digital data

During the transition to wide bandwidth communications channels, data rate reduction will be the driving force as the cost per bit remains relatively high. As the cost per bit declines, the emphasis will shift to scalability. This will be due largely to the market advantages of maintaining a single data file that can be delivered to a wide range of users at different levels of the display hierarchy.

3.4.2 Programmable Decoders

Another major trend that is anticipated is the evolution from fixed single-standard decoders to programmable decoders that can adapt to scalable image representations. Single-standard decoders will be used primarily for devices that tap into the communications network and deal only with one type of image representation. Programmable decoders will deal with families of standards. Fax machines serve as a good example of single and multiple standard encoder/decoders. The Group 1 fax standard provided a single level of resolution: machines were expensive and their use was limited. With the addition of Group 3 fax standards, multiple levels of resolution were supported, including the older Group 1 format. Due to advances in technology, the new machines were better and cheaper, yet compatible with the existing Group 1 machines. The marketplace responded in a very positive manner.

Programmable decoders will be the key component in providing extensibility to the digital imaging architecture. Because of the diversity of image compression standards (Group 3 fax, H.261, JPEG, MPEG, DVI, etc.), these decoders will play an important role in the integration of video and high resolution imaging with desktop computer workstations. This same diversity, with the addition of a digital television standard (or standards) will lead toward the use of programmable decoders in home entertainment and information delivery systems. Essentially fixed solutions will drive the low end of the market, providing inexpensive mass market consumer products, while programmable solutions will dominate at middle and upper levels of the transmission and display hierarchies.

3.4.3 Trends in Display Technology

The use of scanning CRT display technology for certain applications is expected to decline over the next decade as LCD based direct-view displays and projection systems are perfected. LCD displays are used extensively today in portable computers, and LCD light valves for high resolution projection are showing great promise. In a light valve, the LCD is used to control the amount of light – from a flicker-free light source – that can pass through each pixel location: since the display is no longer the light source, significant improvements in brightness can be achieved.

The characteristics of LCD displays are significantly different from flying-spot scanning CRT displays. Flying-spot systems must operate at refresh rates above the critical frequency for flicker fusion: display brightness is limited since the spot is the only source of illumination (most of the display is decaying at any point in time).

Every pixel in an LCD display receives constant illumination. LCDs can be characterized as having long persistence: in fact, a significant design challenge has been to provide faster pixel response to deal with full motion video. This has been accomplished through the use of a transistor at each pixel location (an active matrix display), providing rapid response for pixel replenishment.

The nature of the active matrix circuit also allows a pixel value to be held for at least one second without replenishment, giving the display characteristics similar to a frame buffer. Direct addressing of each pixel location would make it possible to update only those pixels which change from one refresh period to the next. Transmission systems that utilize digital compression techniques to eliminate interframe image redundancies may take advantage of these aspects of LCD displays to implement conditional replenishment.

3.4.4 Conditional Replenishment

A significant portion of the data rate reduction achieved by digital image compression techniques deals with the elimination of interframe redundancies. In essence, much of the complexity, and hence the cost, of these encoding systems involves the processing required to analyze motion image data streams to determine which pixels have changed between temporal samples.

Over the next 10 to 15 years image acquisition and display technologies are likely to move to conditional replenishment. Image acquisition systems may evolve with on-board digital processing to implement conditional image acquisition. These cameras will be programmable.

offering several advantages over scanning cameras that continuously update the entire image raster, including the ability to:

- Acquire high spatial resolution and high temporal resolution data at different rates
- Adjust sensitivity as a function of the temporal update rate
- Output information about multiple objects within the image and their motion vectors
- Relate movement of the sensor (pan and tilt) to image content, allowing for update of only the newly acquired portions of the image

Future display technologies are likely to evolve around direct view displays (possibly LCD) offered in different pixel densities. Direct addressing of LCD displays will allow the use of conditional refreshment of only those pixels that change from one refresh period to the next: the display itself may become the frame buffer, allowing portions of the image to be updated at different temporal rates. Or, combined with an appropriate multi-ported frame buffer design, such a display could support multiple temporal refresh rates simultaneously for different image streams.

4 Critical Issues

4.1 Introduction

The Task Force has identified seven issues that are considered critical to the achievement of the objectives. Many of these issues are, by their nature, complex.

Backward compatibility to existing systems and extensibility to future systems present many technical challenges. The greatest challenge lies in preserving the value of existing infrastructures while enabling an orderly transition to the new architecture. For example, immense investments have been made in the acquisition and transmission infrastructures of our existing NTSC, PAL and SECAM television systems. Likewise, billions of consumers have invested in receivers and video recorders that support these systems. It is equally critical that investment in the vast archives of information and entertainment programming that exist today on film and video be protected, and that the new architecture unlock the economic potential of these archives.

In deliberating on these critical issues, every effort has been made to balance the interests arising from those investments with the future benefit to all of a single global standard. These deliberations have also taken into consideration the installed bases of computer, medical, engineering and scientific imaging systems, and the diverse applications for still imaging in electronic publishing, visual databases and communications. Existing systems that demonstrate interoperability and extensibility – including some which have in fact been extended – were considered. Examples include the French Minitel system and the family of international facsimile standards.

The seven critical issues are:

- The Establishment of Scalable and Interoperable Hierarchies for Basic Image Parameters
- The Establishment of an Appropriate Relationship Between Image Acquisition and Display Refresh Rates
- The Use of Square Sampling Grids (Square Pixels)
- The Establishment of Appropriate Representations for Colorimetry, Dynamic Range and Transfer Characteristics
- The Use of Coherent Image Sampling (Progressive Scanning)
- Identification of the Characteristics of a Digital Image Stream (Headers/Descriptors)
- The Establishment of Appropriate Levels of Compatibility with Current Television and Motion Picture Standards

4.2 The Establishment of Scalable and Interoperable Hierarchies for Basic Image Parameters

An ideal digital image architecture would allow the following image parameters to be independently varied, over a range of appropriate values:

- Spatial resolution
- Field of view
- Viewing distance
- Aspect ratio
- Image acquisition rate
- Display refresh rate
- Dynamic range
- Colorimetry

While this independence may be technically feasible within the fifty year life span desired for the first digital imaging architecture, it does not appear to be practicable for immediate implementation, nor is it required. The choice of an appropriate hierarchy for each these parameters can provide adequate degrees of freedom for system design, while facilitating affordable, high quality transcoding between the levels in each hierarchy.

Scalable and interoperable hierarchies offer many benefits when communications channel issues are considered. Such an approach promotes effective utilization of existing communications channels and the development of new broad band communication services. The lower levels of the hierarchy provide solutions for the capacity constrained channels that exist today. The introduction of new broad band communications services will enable the use of higher data rates to support the improved performance available at higher levels in each hierarchy.

A digital image architecture that provides interoperability across applications with different spatial resolution requirements must be scalable in terms of resolution as discussed in Section 3.3. Interoperability also requires a family of related image acquisition and display rates. The greatest benefit, in terms of cost and simplicity, is gained when the display operates at the same rate as, or an integer multiple of, the image acquisition rate. Though more expensive to implement, the greatest performance benefit is gained when motion compensation techniques are used in encoders/decoders to create in-between frames for display. Section 5.4 discusses the requirements for such a family.

To facilitate this hierarchical approach to a digital image architecture, a scalable approach to image coding is required. Furthermore, improved techniques for video compression are likely to be enabled by the geometric progression in computational hardware. The design of the architecture must make provisions for this progression. Section 5.5 discusses the use of scalable coding algorithms.

4.3 The Establishment of an Appropriate Relationship Between Image Acquisition and Display Refresh Rates

In early discussions about the use of digital codings for HDTV systems, it became clear that receivers would likely need one or more frame stores to implement image decoding. This prompted the idea that image acquisition rates could be decoupled from display refresh rates – the display could be refreshed at a rate that is an integer multiple of the acquisition rate. For this reason the questions of image acquisition rates and display refresh rates will be considered separately.

No topic generated as much discussion in the Task Force as image acquisition and display refresh rates. This is due in part to the diversity of rates that exist in the standards and resulting practices within each of the affected industries. The issue is further complicated by the evolution of television down parallel paths with respect to field rates. Their harmonization will require solutions that lie in the realm of digital technology as well as the realm of politics and negotiation.

The choice of an image acquisition rate is a tradeoff between motion rendition and the resulting data rate. The following considerations are important in establishing a family of acquisition rates.

- The ability to adjust the relationship of motion rendition and data rates to meet specific application requirements
- The cost and quality of transcoding between acquisition rates
- The family should include 24 or 25 fps for acquisition of film – to include both would result in two families of image acquisition rates with poor interoperability
- For North America a family based on 12 fps is one possible choice. However, if agreement were possible on a global standard, a 12.5 fps related family might be acceptable in North America, recognizing that this would require a 4% speed change for acquisition of 24 fps film.

There are many factors affecting the choice of a display refresh rate including:

- Display technology – CRT, LCD, laser, etc.
- Size
- Brightness
- Lag
- Application
- Cost

Refresh rate will be determined by the above criteria and price/performance requirements established by market factors.

Experience has shown that for wide-screen CRT displays of high brightness, a refresh rate in the region of 72 to 75 Hz is required to achieve tolerable levels of wide-area flicker (see Section 3.2.5). In some situations refresh rates in excess of 100 Hz may be desirable. Receivers which operate at 100 Hz (double the normal 50 Hz interlaced scan rate) are being introduced in the 50 Hz market; rate doubling receivers operating at 120 Hz are also being developed for the 60 Hz market.

The relationship of display refresh and image update rates should be based on a progression that permits non-interpolative transformations between the acquisition and display rates in the new architecture (i.e., display at integer multiples of the image update rate). As an example, theatrical display of film is usually double or triple shuttered to minimize wide-area flicker of the display.

Further research into the choice of a single family of acquisition rates and display rates is required. An appropriate interoperable family should include a 24 or 25 fps image acquisition rate which would enable a 72 or 75 Hz display refresh rate. This is the subject of further discussion in Sections 5, 7 and 8.

4.4 The Use of Square Sampling Grids (Square Pixels)

The computer graphics, image processing, and publishing industries have adopted the use of geometrically square pixel sampling grids (frequently simply referred to as square pixels). The use of square pixels facilitates:

- The interchange of image data across standards boundaries
- Development of common storage and transport services across industries
- The use of common equipment, allowing the costs to be amortized across all users

Early on, the computer graphics industry sought ways to insulate applications from variations in display technology. Support for different pixel configurations required run-time transformation of all graphical objects. Even then, applications rarely looked the same from display to display because different pixel configurations caused a

variety of artifacts. These stopgap measures constrained functionality, reduced performance, and added cost to equipment and services. Ultimately, this approach failed.

Instead, computer graphics gravitated towards a common display technology based on square pixels. This simplified system design, which led to lower cost and better performance, enabled equipment and services to be used as commodities across a broad set of industries. Today the computer industry is a major consumer of displays, second only to consumer television receivers.

The use of a common pixel geometry eliminates the need for interpolative resampling when sharing imagery among all users. Resampling has two costs:

- It can be lossy – it can degrade image quality and cannot be reversed without additional loss
- It is costly – it adds a computational step or a hardware equivalent to the process of sharing imagery

Thus the adoption of a common sampling grid is a key issue for discussion and resolution in the SMPTE work towards the specification of a digital image architecture.

4.5 The Establishment of Appropriate Representations for Colorimetry, Dynamic Range and Transfer Characteristics

The concepts of interoperability, scalability and extensibility apply not only to the sampling of the image but equally to the expression of its brightness and colorimetry. A digital image architecture must deal appropriately with dynamic range and colorimetry requirements of the acquisition (including processing), transmission and display modules of a system. During the acquisition and processing of imagery (for example, post-production), image data that may not be required by the human visual system or reproducible on a given display may be required by processing hardware for optimal results. Similarly, the architecture must accommodate image exchanges between systems having differing dynamic range and colorimetry characteristics. The essential issues are summarized in the sub-sections which follow.

4.5.1 Extensibility

Existing image systems can reproduce only a limited range of the colors visible in the real world, often restricted to those corresponding to illuminated objects and the specific needs of the application. The colorimetry of television is currently confined to that of the display device. Figure 4.1 is a color space, within which are illustrated a typical red, green, blue (RGB) gamut of additive primary colors, and a typical yellow, cyan, magenta (YCM) gamut of subtractive colors. Also shown is the hue and saturation representation: saturation is the radial distance from a specified white point; hue is the associated angle. Hue and saturation vectors are shown pointing to the RGB and YCM color gamuts, as well as one vector that extends beyond both. It can be seen that this representation can be used to represent any visible color. This color space is application and device independent.

In the future it may be possible and desirable to extend the colorimetry representation to include a wider range of colors, possibly even including those of self-luminous objects, as one example. A close examination of this issue is needed to establish the range of colors to be represented within the colorimetry of the digital image architecture.

A similar situation to that of colorimetry exists for the representation of dynamic range transfer function. Current systems are individually optimized for the current technology and application and are not easily amenable to an increase in dynamic range. Mechanisms to effectively handle a much wider dynamic range need to be identified.

4.5.2 Scalability

To cover the intended range of application, it is necessary that the color and dynamic range representations be capable of being scaled, preferably independently. For instance, the display of an image having a wide color gamut at the source must produce acceptable color on a display of limited color capability. The reverse situation is also true. Similarly, the display of an image of high dynamic range should not loose essential information when viewed on a display of low dynamic range. The representation must accommodate these requirements efficiently.

The situation is somewhat similar to that of motion picture film in which the latitude of the negative film enables exposure and color adjustment after the image capture, and the *S*-curve of the film characteristic provides effective compression of the highlights and dark regions. Similar provisions may be required in digital image systems to provide reasonable representations for both small and large numbers of bits. A further consideration may concern the optimal distribution of any necessary compression/expansion in respect to overall image quality.

4.5.3 Interoperability

Interoperability demands that the chosen colorimetric representation and the dynamic range representation be device independent for current and future devices. In this fashion, devices supporting differing colorimetries and dynamic ranges can be supported.

It is also important that images of differing colorimetry and dynamic range at the acquisition device should be able to be combined effectively into a single image, when appropriately scaled.

The color space and dynamic range representations that could meet these objectives require extensive consideration. Section 7.7 includes a number of questions that should be considered in the analysis of these and other colorimetry issues.

4.6 The Use of Coherent Image Sampling (Progressive Scanning)

Historically, interlace has been used to achieve a 2:1 reduction in bandwidth requirements (i.e., data rate), and to eliminate wide-area flicker on scanning CRT displays. The use of progressive scanning is nearly universal in computer display applications, and is employed in some high quality video presentations.

In section 3.2.5 it was established that higher scanning rates are required with displays that cover a wider field of view and/or operate at higher levels of brightness than today's television systems. Decoupling the refresh rate of the display from the image update rate provides a mechanism to deal with wide-area flicker – this is discussed in section 4.3.

There are two methods of sampling a moving image that generate coherent image samples:

- Traditional progressive scanning samples the image from top to bottom. The bottom of the picture is sampled later in time than the top of the picture, but in any small area the samples may be regarded as temporally coincident.
- Shuttered devices, such as film and most CCD sensors, expose all image elements simultaneously. Progressive scanning of film by an electronic image sensor thus results in coherent image samples. The use of CCD image sensors which employ electronic shuttering provides a similar mechanism to film; the analog output of the CCD can be sampled progressively as the coherent image data is shifted out of the CCD.

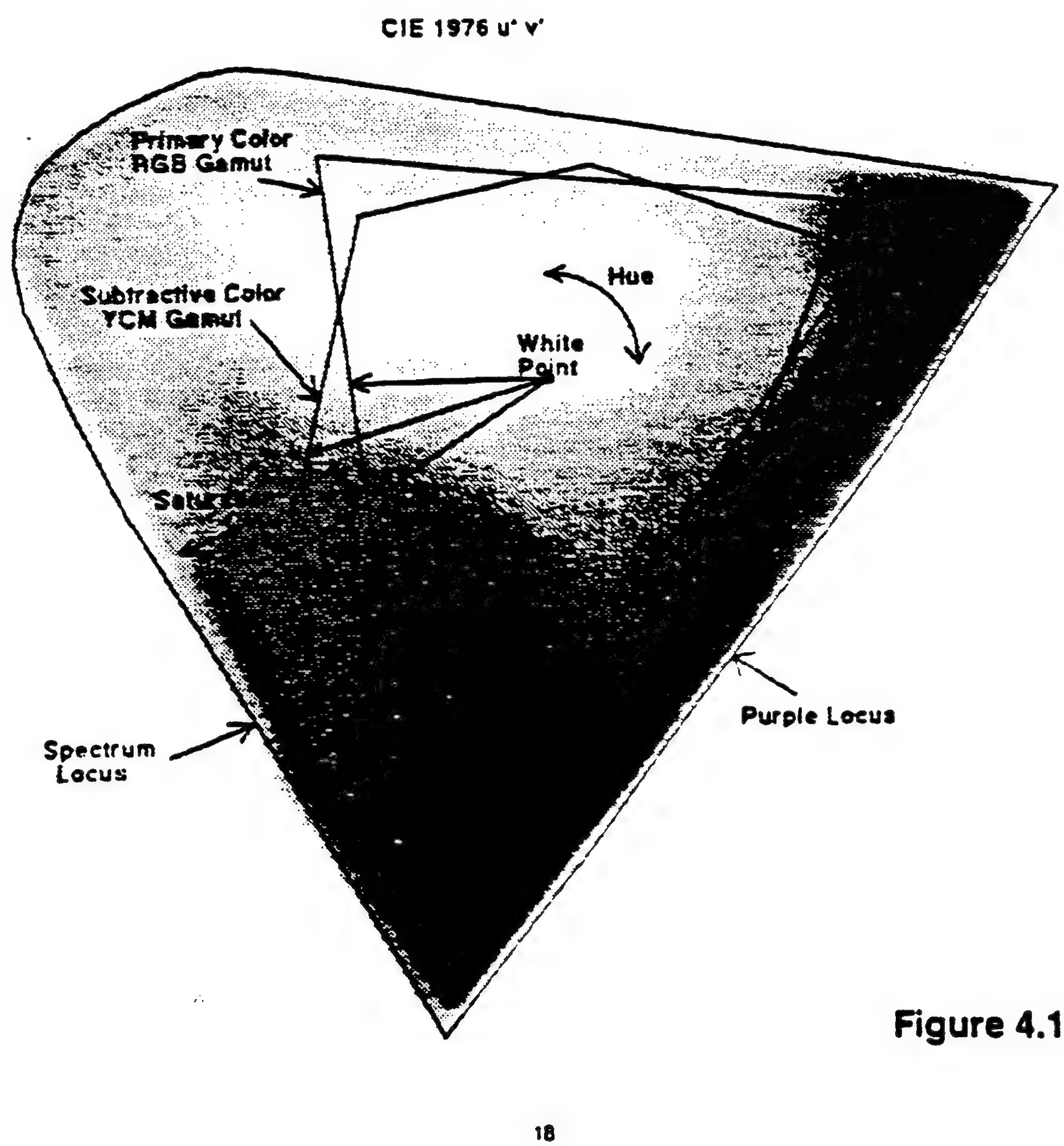


Figure 4.1

Coherent imagery provides the following benefits over interlace:

- Eliminates small-area flicker in high spatial resolution imagery
- Improves the performance of digital video compression and many other processes
- Promotes interoperability by improving all interpolative processes
- Improves the quality of still images generated from individual frames

Fine horizontal lines or edges are commonly used in computer graphics. These objects do not flicker when viewed on a progressively scanned computer display, but may flicker if viewed on an interlaced television display. Many processes – including bandwidth compression, image transformation, and standards conversion – are facilitated by coherent imagery.

The use of coherent image samples simplifies the process of temporal and spatial conversions required for interoperability between different video systems, as well as with other applications and industries. The current practice in standards conversion for interlaced video systems is to first de-interlace the imagery. De-interlacing involves the reconstruction of the missing alternate scan lines; this usually requires the computation of motion vectors.

Interlace provides acceptable motion rendition at lower data rates than progressive scanning, and may improve sensitivity in some camera designs. With the use of appropriate de-interlacing equipment in the studio environment, images acquired with interlaced systems may be converted to progressive scan for subsequent processing. This transformation produces the best results when the interlaced field rate and the progressive frame rate are the same or are related by integer multiples.

Coherent sampling, along with the use of mechanical or electronic shutters, can improve both motion video and still image acquisition, allowing video cameras to be used for a broader range of applications.

The use of progressive scanning is seen by some as a requirement for a digital image architecture, especially when interoperation across applications and industries is a requirement.

4.7 Identification of the Characteristics of a Digital Image Stream (Headers/Descriptors)

A fundamental prerequisite for interoperability in digital systems is a mechanism for identifying and describing digital image data. For this information to be shared, decoders must be capable of identifying and conforming to the incoming data. Even simple decoders – those that only recognize a single standard – must identify data streams which they can decode. This is one of the primary functions of the header. Decoders must also ignore unrecognized data, to allow for extensions to the data stream.

Descriptors provide application oriented information, such as image and coding parameters, processing history, identification of program content, copyright, and scrambling. They also enable extensibility; the descriptor may also contain the coding algorithm or language representation necessary to interpret the encapsulated data. This provides a mechanism whereby expert groups can create and standardize the transmission of messages to meet their needs.

Descriptors may be used to identify and describe data at different levels of an image hierarchy, thus allowing a display system to decode only that part of a stream necessary for its function or capability. Descriptors might also contain information about the preferred display characteristics for imagery.

Thus information such as the colorimetry of the original acquisition system, and the transfer characteristics of the process used to move images from one medium to another, can be included with the data. Decoders would use this information to optimize display of the image.

The SMPTE Task Force on Headers/Descriptors in their Final Report dated January 3, 1992, and approved by the SMPTE Standards Committee on February 6, outlined the criteria for the use of Headers/Descriptors. Work is now progressing on the development of proposed SMPTE Standards, Recommended Practices and Engineering Guidelines.

4.8 Compatibility with Current Television and Motion Picture Standards

The installed base of NTSC, PAL and SECAM equipment within the program production community, together with massive consumer investment in compatible receivers and VCRs, must be supported in the transition to a digital image architecture. Of even greater importance is the requirement to preserve the value of the archives of programs that have been created for mass market distribution using these systems and to exploit these resources to the greatest extent possible in the future.

This is by far the most critical issue of all, so much so that its impact is clear in the discussions of many of the previous issues. Only the last of them, the use of headers/descriptors, is without precedent in existing entertainment industry practice. It is precisely where a dichotomy exists in current practice that the greatest controversy arises – on the issue of temporal rates.

The convergence in being digital may provide the solutions which will resolve the temporal rate issue; convergence around the common language of digital coding, the progression in CPU performance, and the ability to design inexpensive modular interfaces in the form of mass produced microchips.

It is likely that a number of solutions will evolve to facilitate interoperability between the existing world of film and analog television, and the new digital image architecture. These solutions should provide a variety of price/performance options appropriate to the applications requirements.

5 An Example of a Hierarchical Digital Image Architecture

This section suggests a technology transparent hierarchy – one compatible with the present and extensible for the future.

To illustrate the model, specific numbers have been chosen that take advantage of the mathematical relationships discussed in Section 4, as well as the architectures of digital memory and processing components. These numbers are not intended as the basis for a standard, but rather, provide a starting point, from which the validity of the architectural concepts can be verified. Further work is required for verification of the model and determination of the exact numbers upon which a standard can be based (see Section 7).

The following parameters of a hierarchical digital image architecture are discussed in this section:

- Open Architecture
- Multiple Spatial Resolutions
- Image Acquisition Rates and Display Refresh Rates
- Scalable Coding Algorithms

5.1 Open Architecture

In Section 3 it was indicated that the opportunity exists to design an open digital image architecture based on generic, inexpensive, and increasingly powerful digital components.

For a digital image architecture to be cast as an open system, two steps are required:

- Modularization into acquisition, transmission and display
- Defining standard interfaces between these modules

There must be a systems engineering of the standards so that the modules work together. There are two basic interface definitions to be publicly standardized:

- Headers/descriptors – a data definition interface between the acquisition and display modules. The headers/descriptors issue (Section 4.7) attempts to address how the acquisition and display segments interact.
- Communications service definition – this interface defines how the acquisition and display modules state communications requirements to the transmission module. This is the vehicle for an end sub-system to state its communications requirements (latency, bit rate, error budget, synchronization, interactivity, etc.) to the transmission system and for the transmission sub-system to inform the display sub-system of these parameters.

Some of the parameters that should be part of this communications service definition include:

- Error tolerance – some data types and some applications require bit-perfect delivery. Headers/descriptors and database transactions are obvious examples. On the other hand, some video and audio applications may be able to stand some noise – how much being a human perceptual, engineering and economic tradeoff.
- Delivery guarantees – some applications require reliable delivery, where lost packets are retried until delivery is acknowledged. Others, such as live viewing or remote display, require that lost packets not be recovered.
- Latency – the amount of tolerable time between acquisition and display, in an active system
- Synchronized service requirement – multiple types of data that need to be delivered together

This careful modularization encapsulates other issues, including the critical issues discussed in Section 4, so that they can be addressed one by one.

It can be argued that there is no need for rigid architectural standards in a digital world: that programmability in the transmission and display hierarchies provides a sufficient basis for interoperability. Perhaps some day this will be true. If the goal of longevity for the first digital image architecture is achieved, it is likely that the designers of the next imaging architecture will be less constrained than we are today.

The first digital image architecture, however, must provide a bridge from the closed systems of the past to the open systems of the future. The fundamental structure of the digital building blocks and economies of scale associated with standardization suggest that the organizations charged with establishing these standards work in harmony.

5.2 Designing Display Systems to Deal with Multiple Spatial Resolution Requirements

The perceived resolution of a display is determined primarily by the viewing distance and the visual acuity of the observer. Visual acuity is often determined using sets of alternating black and white lines of equal width. One black/white line pair represents one cycle. The number of cycles that can be resolved across one degree of the eye's viewing field is typically used as a measure of human visual acuity, and is stated in cycles (line pairs) per degree. Under some conditions, with high contrast line pairs, human visual acuity extends beyond 40 cycles per

degree: approximately 22 cycles per degree is perceived as a sharp image.

If the resolution of a display is held constant and the viewing distance is a variable, the resolution perceived by the viewer – measured in cycles per degree – will increase as the viewer moves away from the display. Therefore, all displays can be considered to be high resolution if viewed from an appropriate distance.

At a distance that varies with the visual acuity of each individual, the actual resolution of the display equals the limit of that viewer's ability to resolve image detail. Beyond this viewing distance additional image detail cannot be perceived: that is, the display has more resolution than is required for this viewer and set of viewing conditions.

In some cases excess resolution may be desirable. For example, the operator of a personal computer can typically reduce the viewing distance to a high resolution desktop display by one-half, simply by leaning forward, thus taking advantage of additional resolution. In this example perceived resolution improves enough to be significant, while moving 15 inches in a movie theatre would have little effect on perceived resolution.

The NTSC transmission standard was designed to provide a resolution of approximately 21 cycles per degree over a viewing field of just under 11 degrees. Display size can be a variable in today's television, ranging from a diagonal of a few inches (a personal display) to more than 30 feet (direct view displays in stadiums and projection displays in controlled lighting environments). These displays differ only in the size of their pixels. At the appropriate viewing distances, the perceived resolution of the personal display and the stadium display will equal the design goal of 21 cycles per degree, and both displays will cover 11 degrees of the observer's field of view.

Many display applications require higher levels of perceived resolution. To increase the level of perceived resolution, while holding viewing distance constant, additional samples of the same image must be added, increasing pixel density. To cover a wider field of view, as in wide-screen displays, holding the same viewing distance and perceived resolution, new information, at the same pixel density, must be added to extend the picture.

5.3 Defining a Spatial Resolution Hierarchy

Section 3.2 identified the need for a variety of image resolutions to deal with specific imaging requirements. These ranges can now be further defined in terms of field of view and resolution in cycles per degree.

With personal, home entertainment and theatre displays, the viewer can vary the distance from the display, and thus vary the perceived resolution, over a significant range (see Figure 5.1). Taking into account the variations in acuity in the population, and variations in viewing distance for each application, it is common practice to design a display system for the average viewing conditions in each application. The overlaps in cycles per degree between low, normal and high resolutions are shown in the table to account for these variations.

Resolution	Cycles per Degree
Low	1 - 15
Normal	10 - 25
High	20 - 30
Ultra High	30 - 40

A special case exists for head mounted displays which provide a fixed viewing distance: here the display manufacturer must select the level of resolution appropriate for the application and then design for a specific perceived resolution.

Using these guidelines, a high resolution display designed for a 35 degree field of view would require about two thousand pixels per line at 30 cycles per degree. In a desktop computing application where the viewer is 30 inches from the display, the length of an active line (display width) would be about 19 inches. In an entertainment application, such as a consumer television receiver viewed from a distance of 108 inches (9 feet), the length of an active line would be about 68 inches.

These examples are illustrated in Figure 5.2. In this figure the principles described in this section are used to illustrate the relationships between the four resolution levels of the model hierarchy and a variety of display applications. The numbers, especially as they relate to image size (in pixels) are entirely relative; they serve only as examples of the pixel count required, at average viewing distances and fields of view, to achieve the specified perceived resolution.

It is important to note that seemingly diverse applications such as personal computer and home entertainment displays will have similar resolution requirements as the size of the home entertainment display increases beyond the narrow field of view of today's television receivers. It is also important to note that direct view CRT displays (which are currently limited to around 40 inch diagonals) require resolution in the normal range for home entertainment applications.

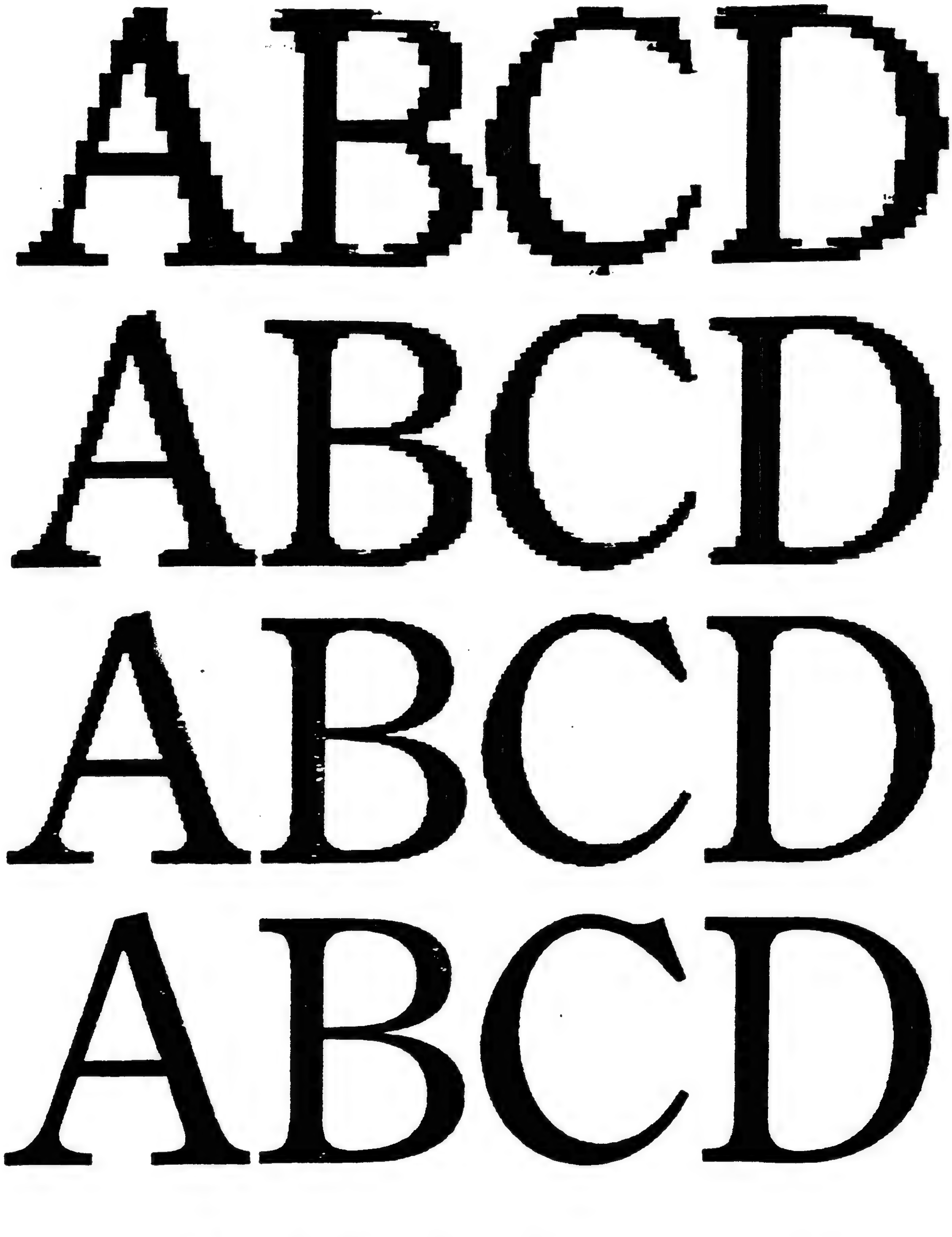
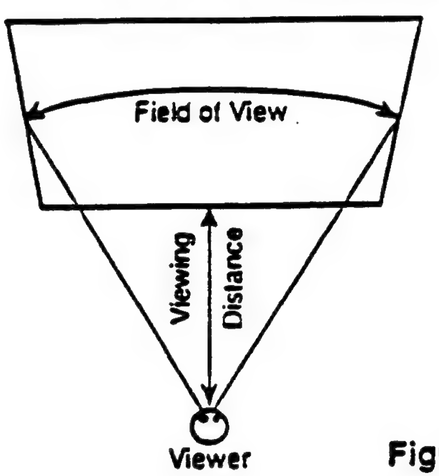


Figure 5.1 - Relative Tile Resolutions - These groups of letters represent the relative resolution for each level of the hierarchy from Level 1 (top) to Level 4. To better understand the practical application in displays, place this figure where it can be viewed from a distance of between 30 inches and 15 feet. Level 4 should be *sharp* at 30 inches; as you move away each level lower in the hierarchy will become sharp.

Resolution Level	Average Field of View	Average Viewing Distance	Target Display H x V (rounded)	Pixels H x V (square)	Approximate Resolution cycles/degree
Level 1 Low Resolution	25°	30" 76 cm	12" x 7" 30 x 18 cm	512 x 288	10
		108" 274 cm	44" x 25" 112 x 64 cm		
Level 2 Normal Resolution	25°	30" 76 cm	12" x 7" 30 x 18 cm	1024 x 576	20
		108" 274 cm	44" x 25" 112 x 64 cm		
Level 3 High Resolution	35°	30" 76 cm	19" x 11" 48 x 28 cm	2048 x 1152	30
		108" 274 cm	68" x 38" 173 x 97 cm		
		360" 914 cm	227" x 127" 577 x 323 cm		
	50°	108" 274 cm	101" x 57" 257 x 145 cm		20
		360" 914 cm	336" x 189" 853 x 480 cm		
Level 4 Ultra High Resolution	50°	30" 76 cm	28" x 16" 71 x 41 cm	4096 x 2304	40
		108" 274 cm	101" x 57" 257 x 145 cm		
		360" 914 cm	336" x 189" 853 x 480 cm		



AVERAGE VIEWING DISTANCE	
Distance	Applications
30" 76 cm	Personal and Computer Displays
108" 274 cm	Home Entertainment Systems
360" 914 cm	Theatrical Display and Business Presentation

Figure 5.2 – Example 16 x 9 Aspect Ratio Displays

5.3.1 Key Concepts of the Model

The example spatial resolution hierarchy is designed around a few basic concepts:

- An integer progression of hierarchy levels based on commonly used digital processing and memory architectures
- Support for a variety of aspect ratios and spatial resolutions based on the concept of image tiles
- The ability to construct displays for virtually any application requirement from tiles of the appropriate resolution

The hierarchy progression is based on the use of integer values related by powers of two. Essentially, at each higher level of the hierarchy, resolution doubles (e.g., 1, 2, 4, 8, etc.); subsets of the lowest level can be derived similarly (1/2, 1/4, 1/8, etc.).

It is noteworthy that such sequences also appear in the computer processor and memory component industry. This approach takes full advantage of the generic building blocks that are the driving force in the transition to a digital world.

In order to provide continuity between the various resolution levels of the hierarchy, the model is based on the concept of an image tile. For the purposed of this discussion, a tile can be considered to be a constant portion of an image, representing the same part of the image regardless of the resolution level or image size. Thus, at each higher level in the hierarchy, the resolution within a tile doubles in each axis. This is illustrated in Figure 5.3.

The power of two progression may now be applied to determine the resolution, in pixels, for each level in the hierarchy.

Level	Name	Resolution in Cycles per Degree	Pixels in One Tile	Pixels in 32 x 32 Tile Superset
1	Low	1-15	16 x 16	512 x 512
2	Normal	10-25	32 x 32	1024 x 1024
3	High	20-30	64 x 64	2048 x 2048
4	Ultra High	30-40	128 x 128	4096 x 4096

In this model a tile represents an area equal to 1/32nd of the image at any level of the hierarchy. Thus each level consists of a 32 x 32 set of tiles (see Figure 5.3). The selection of this fraction for a tile is arbitrary; it was chosen because it is a convenient building block – integer multiples can be used to construct displays at all of the aspect ratios and spatial resolutions discussed in the model.

5.3.2 Construction of Displays from Tiles of the Appropriate Resolution

The table in Figure 5.3 provides a matrix of display aspect ratios and resolutions that can be derived from the full set of 32 x 32 tiles at each level. Since the tile size is a constant, each column represents a constant size display at four perceived levels of resolution.

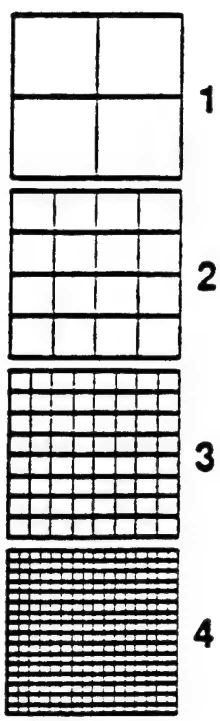
The diagram in Figure 5.3 establishes several important relationships that provide a bridge to the past and illustrate how interoperability can be achieved:

- Region A (4:3) shows the relative spatial resolution of 525 line systems.
- Region B (4:3) show the relative spatial resolution of 625 line systems.
- Region C (16:9) is representative of the display that some industry observers expect will be used in the next generation of home entertainment systems.
- Region D (4:3) is commonly used in computer and entertainment systems. It allows the incorporation of regions A, B, and C, which correspond to traditional video and film aspect ratios, along with additional areas for the display of Graphical User Interface (GUI) elements. It is important to note that the 1024 x 768 and 640 x 480 resolutions – shown for level 2 – are both popular display resolutions for personal computer systems.
- Region E (1:1) is emerging as a new display aspect ratio for certain applications. The FAA recently began procurement of 2048 x 2048 square pixel, progressive scan CRT displays for air traffic control systems.

The tile concept can similarly be applied to the manufacture of displays. In this case, a physical display tile would correspond to a conceptual tile and would have different physical sizes for different size displays and different pixel densities for different resolution requirements. Similarly, displays of different aspect ratios could be constructed by the selection of the appropriate conceptual tiles as shown in Figure 5.3.

Thus, using tiles and only four resolution levels, it is possible to construct a display for virtually every possible application; furthermore this display can also be used to show imagery from other levels of the hierarchy. This is especially practical if a scalable coding architecture is implemented that conforms to the same resolution progression.

TILE RESOLUTION IN PIXELS Relative tile resolutions are shown below		NUMBER OF TILES IN DISPLAY 20 X 15 4:3 Subset Region A	24 X 18 4:3 Subset Region B	32 X 18 16:9 Subset Region C	32 X 24 4:3 Region D	32 X 32 1:1 Region E
LEVEL 1 Low Resolution	H = 16 X V = 16	H = 320 X V = 240	H = 384 X V = 288	H = 512 X V = 288	H = 512 X V = 384	H = 512 X V = 512
LEVEL 2 Normal Resolution	H = 32 X V = 32	H = 640 X V = 480	H = 768 X V = 576	H = 1024 X V = 576	H = 1024 X V = 768	H = 1024 X V = 1024
LEVEL 3 High Resolution	H = 64 X V = 64	H = 1280 X V = 960	H = 1536 X V = 1152	H = 2048 X V = 1152	H = 2048 X V = 1536	H = 2048 X V = 2048
LEVEL 4 Ultra High Resolution	H = 128 X V = 128	H = 2560 X V = 1920	H = 3072 X V = 2304	H = 4096 X V = 2304	H = 4096 X V = 3072	H = 4096 X V = 4096



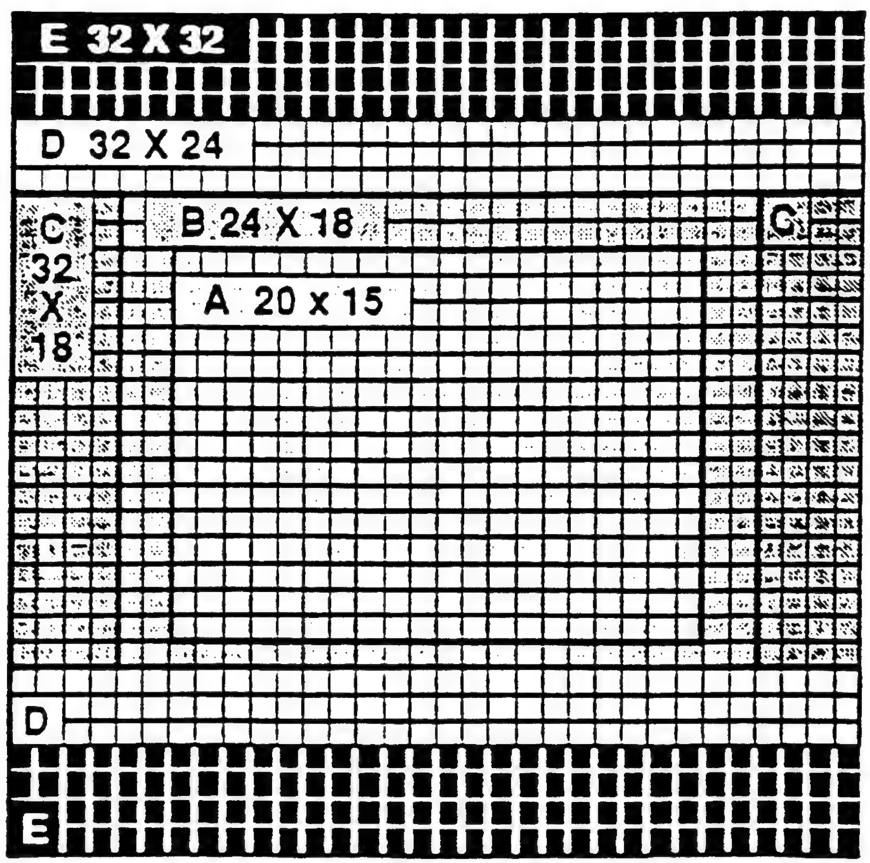


Figure 5.3 – Display Constructions Using Multiple Tiles

5.4 A Family of Related Image Acquisition Rates and Display Refresh Rates

A family of image acquisition and display refresh rates should be based on a progression that permits non-interpolative transformations between the acquisition and display rates. This is easily implemented if the acquisition and display rates are the same, or if the display refresh rate is an integer multiple (or fraction) of the image acquisition rate.

Since significant archives of high resolution program material exist on film, which was acquired at 24 or 25 fps, one of these rates should be included in the progression. A progression based on integer multiples of 12 would include 12, 24, 36, 48, 60, 72, 96, 120 Hz, etc. A progression based on integer multiples of 12.5 would include 12.5, 25, 50, 75, 100, 125 Hz, etc. These progressions might also include integer fractions of 12 or 12.5 (e.g., 1/2 or 1/4 of the base frame rate for applications such as videoconferencing and searching of video databases.)

It has been common practice in Europe to display 24 fps film at 25 fps for compatibility with PAL and SECAM: this results in a 4% speed increase. Many European programs produced for television distribution are acquired at 25 fps: if the family of rates is based on 24 fps, these programs would be played 4% slower. As indicated in Section 4.8, further research is required to determine the impact of choosing one of these rates on those industries that utilize film for image acquisition.

Ideally, compatibility with existing electronic imaging systems would be accommodated in the design of the standard modules that will interface these systems with the digital image architecture. By design, this would place the burden of compatibility on the systems that are being replaced rather than products that conform to the new architecture: thus the future will not be constrained by today's limitations.

In the process of developing the existing analog and digital high resolution television systems, the designers of these systems have demonstrated the practicality of such a modular approach to interoperability. A variety of translation devices have been demonstrated that allow interoperation between PAL, NTSC, HD-MAC and MUSE. The interface modules that will be required to transform the signals from these systems (especially NTSC and PAL) into the new architecture offer the potential for large volumes. It is likely that the market for these modules will be characterized by intense competition, leading to a range of solutions at various price/performance levels.

In the near term the choice of a family of rates, based on 12 or 12.5 Hz, would provide optimally low cost and high performance, for both advanced television and computer uses, as well as providing global interoperability. In the longer term, decoupling of acquisition, transmission and display is likely to lead to entirely new approaches to pixel replenishment that may render the current concept of image acquisition rates and display refresh rates meaningless.

5.5 Scalable Coding Algorithms

Scalable image decomposition offers the ability to produce image data in packages that can be combined to produce images at a variety of spatial and temporal resolutions. Decoding and displaying the lowest frequency image packets would produce an image at the first level of the hierarchy. Additional packets (encoded with spatial and temporal differences) would be decoded to produce images at higher levels of the hierarchy.

This approach enables extensibility. For example, the coding of low resolution imagery might remain unchanged to provide compatibility with existing decoders, while new coding methods, made possible by the geometric progression in computational hardware, can be introduced to support more advanced imagery. Increasingly powerful (and affordable) programmable decoders can provide compatibility with the standards that form the foundation of the digital image architecture, and the additional processing power required for future enhancements to the architecture.

6 Industries and Applications Considered

6.1 Industries and Applications

Industries are categorized by current market segments. It is important to keep in mind that convergences among existing industries will likely occur (e.g., computers and consumer electronics; audio, video, and datacomm), and as new opportunities to provide value products and services emerge, entirely new industry segments will undoubtedly come forth.

It is becoming difficult to draw the line, even today, between consumer electronics and computers. Today's video game machines, already in millions of homes, are marketed as consumer accessories to televisions, but are in fact, more computationally competent than personal computers of only a few years ago. Similarly, personal computers are being marketed to the home market through traditional consumer electronics channels.

Traditional business factors should always be considered. These include equipment replacement costs, amortization, benefits, competition, market needs, and access to material.

Successful industry participants will pay close attention to emerging trends and will help to bring them about. Sometimes, deep pockets may be required to create a market. (It took years of major losses in both equipment and programming efforts before color television became profitable.) In contrast, agreement on a common architecture across a wide range of industries and applications would spread the costs and encourage early adoption.

The groupings used for this Report help to relate application requirements to industries. It is well understood that there is already much overlap between industry groups and applications.

The industry groupings are as follows:

- Entertainment Providers (6.1.1)
- Distribution and Communications (6.1.2)
- Commercial Equipment Manufacturers (6.1.3)
- Consumer Electronics Manufacturers (6.1.4)
- Computers and Information 6.1.5)
- Education (6.1.6)
- Engineering and Science (6.1.7)
- Healthcare (6.1.8)
- Military and Aeronautics (6.1.9)

6.1.1 Entertainment Providers

Entertainment provider fields include programming, animation, games (personal and arcade), broadcasting, cinematography, post-production, theatrical presentation, and pre-recorded media.

The technologies used in these fields are highly dependent on downstream profits. It can be difficult to justify large investments (e.g., an HDTV production facility) in new technologies that can only be utilized by a small portion of their market. Smaller investments that require minimal infrastructure changes (e.g., MTS stereo, VHS-HQ) can be more easily justified, particularly when end-users can benefit with existing equipment or rapid upgrade is anticipated. Backward compatibility and extensibility are key issues here and can only be successfully violated when there are substantive benefits to the end user (e.g., audio compact disc).

Revenue streams can often be anticipated to flow well beyond the initial release of the product. Residuals from syndication, rentals, and sales require that providers anticipate future trends in end-user viewing equipment capabilities. This is one reason why most prime time television is shot on 35mm film and not video.

6.1.2 Distribution and Communications

The distribution and communication industries that will be affected by digital image systems include telephone, television broadcasting and cable TV, utilities, video conferencing, electronic mail (including text, data, image, animation, video, and sound), and mobile communications. Carrier channels that will play a yet-to-be-determined role in this process include optical fibers, broad and spot-beam satellites, microwave, cellular, conventional VHF/UHF terrestrial broadcast, broadband coaxial cable, and local and wide area networks. Also impacted will be video tape, video disc, game, and general software distribution.

There is some effort to establish a video dialtone similar, in concept, to today's voice telephone dialtone. As communication networks increase bandwidth, and compression technologies improve, an increased use of remote real-time visual communications can be expected.

These same advancements also facilitate rapid downloading of video information from media servers: at a 100:1 compression ratio, the data for a typical motion picture could be transmitted in a few minutes over a full video-bandwidth capable network.

Because of the universal proliferation and conversion standards for the telephone, it is likely that we will soon see extensions of current fax standards including: voice

fax (voice mail), high resolution color image fax, and video fax (video mail). One of the driving forces behind the development of the JPEG image compression standard was the need for an efficient data reduction technique for the transmission of still images.

The telecommunications industry is well down the road in the establishment of digital imaging standards. The CCITT, which controls fax standards, worked with the IEEE on the JPEG standard and on the videoconferencing standard known as P.64 or H.261. These groups are also responsible for the MPEG family of moving picture standards. JPEG and MPEG I and P.64 form the basis for the first generation of image telecommunications products that are already starting to reach the market.

These standards were designed with a high degree of flexibility to deal with a variety of imaging applications; they have served as excellent examples for the Task Force in the area of interoperability and scalability. Currently the MPEG group is working on extensibility; MPEG II is targeted for the delivery of higher quality moving image data streams in the range from two to forty megabits per second. The MPEG working group is investigating scalability as a requirement for this extension of MPEG. It would be beneficial for these new standards to relate harmoniously to other digital imaging architectures.

The merging of both broadcast and interactive voice, image (including graphics and video), text, and data across diverse transport media will create challenges in properly matching the information with the delivery mechanism. Current efforts to implement interactive television, for example, use differing transmission media for each direction (e.g., broadband in; telephone or cellular radio out).

Factors such as existing infrastructure, projected time and cost to deploy, bandwidth cost, regulatory issues, nature of the signal, target viewer, compression, error sources, localization, security, latency, etc., need to be considered.

The communications infrastructure deployed for the entertainment market could provide a profound leverage for the information domain. For example, a broad consumer demand for access to high bandwidth entertainment (and other) services could accelerate the national installation of fiber-optic cables. Once in place, these high bandwidth networks could also be used as high performance links to super-computers and very large data bases, and broadly distribute real-time business, engineering, and scientific data.

While installation of fiber-optic cable to a major user base can take many years, new or existing satellites can cover huge population areas very quickly. A variation of direct broadcast satellite (DBS) transmission is spot-beam satellite technology. In this approach, as few as three satellites could be used to provide localized high quality (HDTV) signals to small inexpensive receiving devices in as many as 150 geographic areas within a country the size of the continental United States.

6.1.3 Professional Equipment Manufacturers

Equipment manufacturers who produce studio, production, storage and distribution, and test and measurement equipment will enjoy opportunities to provide their customers with new products and services that can be useful across a range of industries. Digital, extensible, scalable image architectures can provide high value per dollar and increased economies of scale.

The computer, medical, and graphics industries could similarly benefit from harmonious formats that would allow them to produce image generating, manipulating, managing, storing, and viewing applications and devices at reduced cost and increased interoperability.

Some specific industrial application areas include security equipment for surveillance and identification, and product and process inspection.

6.1.4 Consumer Electronics Manufacturers

The introduction of digital technologies into consumer products opens the way to new and improved services and capabilities. As the consumer market increasingly demands higher image quality for both work at home (e.g., personal computers) and entertainment (e.g., television and video games), there will continue to be incentives to push the technologies that will bring a better picture to the consumer.

This will create opportunities in the receiving devices, the electronic components that go into them (e.g., semiconductors, light sources and modulators) and the subsystems (e.g., displays, tuners, and signal processors). The likely emergence of new product categories can both heighten and personalize the entertainment experience.

Ancillary devices (e.g., tape and disc recorder/players, camcorders, editing, processing, sound systems, printers, scanners, interactive peripherals) will be additional sources of added value products.

It is likely that computer control technologies will play an ever increasing role in home entertainment and information systems. The integration of all of the equipment listed in the preceding paragraph in the home entertainment environment has proven to be a major problem – and a significant opportunity. We have seen programmable remote control devices evolve to replace the profusion of separate infrared controllers (TV tuner, cable tuner, VCR, laserdisc, audio CD, radio tuner, etc.). The integration of the graphical user interface from the world of desktop computing with the home entertainment/information system has begun.

Collaborative cross-industry efforts will merge computers into home entertainment networks, dealing with the issues of component integration, connection to multiple sources of entertainment and information, user interface, and "user friendly" programming of the system. Various flavors of "personal computers" in the home will be able to connect to this network as well as intelligent appliances and remote control devices. Inexpensive networkable cameras will allow remote visual monitoring: the front door, the baby's room, etc.

6.1.5 Computers and Information

Human vision provides the highest bandwidth information interface to the machine world. Computer technology can serve as an effective enabling tool for image information creation, capture, processing, storage and archiving, access, transmission, and presentation. While computer assisted information in the 1980s largely focused on text, data, and simple graphics, rapid changes are taking place to support other media (audio, image, animation, video, simulations, etc.). This places increased demands on computer performance and human interface to handle the significantly higher data content in these media.

To provide specific types of information to users, new classes of specially tuned information appliances will likely develop. These appliances will rely on information providers to collect, generate, and organize information. In the education market, for example, an information appliance might be tuned toward providing everything a student needs to progress through a particular class. Besides basic course content, texts, lecture notes, assignments, etc., it could make extensive use of imagery to provide interactive multimedia tutorials, remedial help, lab simulations, extensive reference material, electronic messaging, and smart links to classmates.

In the information age, a critical challenge is the productive management of the overwhelming amount of information produced each year. Unfortunately, images and video tend to make this problem even greater. While database search engines deal reasonably well with keyword searches and inverted indexes on textual data, corresponding tools for other media have tremendous opportunities for improvement.

Museums and libraries could use electronic file systems to catalog and view very high resolution images of the masters. Sculptures and other three dimensional objects could be shown on stereographic or holographic displays, or printed on very high quality large format printers.

The role of the artist and graphics designer has changed dramatically as the quality and flexibility of the "electronic canvas" has come to emulate the various forms of traditional media. Just as the camcorder has allowed many budding cinematographers to explore their art, high resolution drawing tools with interactive training are revolutionizing electronic publishing and winning over graphic artists. Many artists are expanding into new markets such as videographics and animation from this electronic base.

Traditional forms of printing and publishing information delivery will continue to exist alongside newer mediums. Electronic billboards could change messages by day of week or time of day. Electronic books, magazines, catalogs, and advertisements can integrate interactive video and other media to tell a story, make a point, or sell a product. They can also elicit information from the user that can provide useful information to the publisher (e.g., "hard to understand this concept," "would like product in green").

6.1.6 Education

One strategy for promoting the use of digital image technologies in education is to leverage high volume consumer products. There is now a real opportunity to leverage scalable, interoperable, extensible consumer products into the classroom and other learning environments (e.g., lab, home, library, tutoring, group study).

Institutional training represents the high end of the educational market. An economic return on investment can often justify the use of expensive technology to maximize training "productivity" since the employee students are being paid wages while not working. Increased use of sophisticated interactive multimedia tools developed and used in these environments could find derivative use in public classrooms and the home.

6.1.7 Engineering and Science

Engineers and scientists have traditionally used the high end of graphics and imaging systems for data visualization, design, simulation and scientific visualization. This will likely continue as new uses expand into such areas as microscopy and astronomy.

This community has often utilized high-end versions of consumer technologies (e.g., TV CRT/Workstation CRT). Their role in leading versus leveraging the next generation of imaging systems is not clear. The existence of a proper digital image architecture will reduce barriers across applications, platforms, and markets.

6.1.8 Healthcare

Healthcare represents a growing cost concern for most industrialized societies. While a digital image architecture may not directly reduce costs, the judicious use of images and video can provide an improved cost/benefit ratio for physician training, medical research, and general patient care.

High resolution imaging can be useful in radiology, microscopy, patient monitoring (especially during surgery), and consultation with specialists in a remote location.

Image requirements can be very stringent. Doctors sometimes use a magnifying glass to look for subtle changes in gray level on an X-ray. Image fidelity is critical.

Training simulators, perhaps utilizing virtual reality techniques, can provide medical students with improved environments for learning over classroom and cadaver procedures.

Although the spatial resolutions and signal integrity requirements may exceed many other applications, the healthcare community would like to benefit from harmonization with other digital image architectures.

6.1.9 Military and Aerospace

Traditionally, the military and aerospace industries have driven the high end of the imaging market with severe mission-critical requirements. While in the past cost was a concern second to functionality, new economics dictate that more effective leverage be made of existing standards, technologies, and products wherever possible.

Typical applications include radar and other tracking, surveillance, flight simulation, general training simulators, mission/situation control rooms, instrument control panels in aircraft/vehicles, satellite imaging, virtual reality, telepresence, and cartography.

Increasing emphasis is being placed on dual use technologies. The community learning network is one example of using advanced imaging technologies for both government and civilian education.

6.2 Application Requirements

Throughout the digital image path, there are specific and interrelated requirements that should be understood for any particular application or family of applications. A standards architect or application developer should be aware of not only current needs and projected needs, but also past infrastructures, potentially related application areas, and long term technology trends.

6.2.1 Latency

For any application, tolerance to transmission path latency should be considered. Small latency times can impose increased compression and transport costs. Some examples of acceptable latencies follow:

< 0.1 second (interactive flight simulation, virtual reality, telepresence, video games, cursor control)

< 0.2 seconds (teleconferencing, video phone)

< 1 second (gambling events, market data)

< 1 minute (live events, news events, announcements, data files)

< 1 hour (weather, general news, images)

< 1 day (newspapers, recorded programs, phone listings)

< 1 week (catalogs, general reference material, maps, business directory pages)

6.2.2 Synchronization with Other Media

Traditionally, motion pictures and video have been concerned primarily with synchronization between image and sound track. Future imaging architectures should consider synchronization requirements with other media and general control inputs and outputs. Examples of other media used for special applications might involve any of the other three human senses (touch, smell, taste), extensions of the first two (sight and sound), as well as physiological inputs (e.g., EKG, EEG, EMG, respiration, perspiration, salivation, biochemical levels).

Both input (i.e., response time) and output synchronization should be considered. Acceptable synchronization can vary with image content. For example, voice should have excellent synchronization when an actor's lips are seen, but less synchronization is needed when the actor is off camera. Background music can accommodate even less synchronization as long as it is not keyed to action or scene transitions.

Motion inputs (e.g., physical controls, gestures, head or eye tracking, facial expressions), and outputs (vibration, g-forces, wind) can also have varying needs for synchronization.

6.2.3 The Digital Image Path

Images flow through (and may be stored in) the following five processes:

- Capture/acquisition/creation
- Processing
- Transport
- Reconstruction
- Presentation

Each of these represents a range of opportunities for industries and applications.

Image capture/acquisition/creation includes:

- Live image capture (video cameras and film cameras)
- Previously recorded motion images (telecine)
- Previously recorded still images (fax, scanners)

- Artificial images created on graphics computer workstations and animation computers

Factors affecting image capture include available light, subject motion, spatial resolution, colorimetry, and cost.

Good light sensitivity is an important factor in available-light location shots. In studio productions, sensitive cameras can reduce equipment and electrical requirements for lighting and resultant air conditioning.

In scenes with high subject motion like sporting events, an image sensor configured with quick response to motion is important. A fast scan rate is the overriding factor here, particularly when minimal single frame blur is important (for slow motion or single frame playback). Current technology favors interlaced image sensors for this type of application, however post-processing or future technical advances in image sensors can be applied to eliminate interlace artifacts before the signal goes very far down the image path.

The scan rate of an image sensor should also relate compatibly to the frame rate of its source material (e.g., movie film), and/or the anticipated frame rate of the viewing device.

Spatial resolution can be expected to improve for both still and motion image sensors. As described elsewhere in this Report, square grids and properly scalable array geometries are important factors in providing extensibility.

Specific applications can require high spatial but lower temporal resolutions. Image scanners fall into this category. Used for medical X-rays, hard copy scanning, film conversion, and fax, a common characteristic is the need for high image integrity (e.g., error-free image sensors, lossless compression, minimal geometric distortion, robust error correction).

An ideal image sensor would be able to resolve the entire range of color tints and hues visible to a human eye over a very wide dynamic range. It would also have well defined electrical transfer characteristics. Falling short of this, it is important that the colorimetric transfer characteristics be sufficiently defined to accommodate faithful propagation throughout the image path.

Future image sensors will likely contain increasing amounts of on-device signal processing in the form of motion detection, compression, and error detection and correction. They may not be scanned, but interrupt driven, responding to changes in the image. Devices may even begin to take on some functional characteristics of the human retina.

Processing includes:

- Format conversion
- On-line and off-line editing
- Color and density control
- Compositing
- Image and sound enhancement and synchronization
- Special effects (blowups, mirror imaging, zooms, pans, superimposition, morphing, titling, speed change, freeze frames, dissolves, wipes, paint box, etc.)
- Storage and retrieval (clips, scenes, shows, stock footage, libraries)

Transport includes:

- Compression
- Encoding
- Transmission
- Movement through air, space, fiber, wire, or highways
- Storage
- Reception
- Selection

Reconstruction includes:

- Decoding
- Error correction and concealment
- Decompression
- Synchronization
- Image enhancement

A receiving device requires an image processing engine to properly reconstruct information from the signal. This information must be compatible with both the presentation device and local storage (e.g., tape, disc, semiconductor).

Presentation includes:

- Compositing
- Overlaying
- Buffering
- Storing
- Manipulating
- Displaying

Presentation manipulation can be spatial (e.g., zoom, pan, detailing, colorization) or temporal (slow motion, still frame, fast scan). Many of these manipulations may

be difficult to achieve with compression schemes that use incremental transmission or sub-sampling techniques.

Some applications require bi-directional capabilities. Some examples are: interactive communications. on-demand programming. pay-per-view. and client/server models.

In a client/server structure. the presentation "client" device may be physically separate from the reception/reconstruction "server." This model might apply to both robust and upgradable servers in a home neighborhood or on a computer network within an engineering office environment. In either case. the server would need to be able to interrogate the client so that it could properly reconstruct the presentation information.

6.3 Displays

6.3.1 General Considerations

No other component of a digital image system has more impact on industries and applications than the display. More than bandwidth limitations. image capture. and signal processing. the performance and economic constraints of available displays are currently the greatest pacing factors.

Applications that can live within current display constraints. or rapidly utilize or promote advancements in both flat screen and projection (and to a lesser degree. direct view CRTs) will be in the best positions to prosper.

Potential display image sizes can range from a wrist watch to a planetarium. General factors that should be considered in specifying a display include: number of viewers. viewing conditions. spatial and temporal resolutions. pixel size and shape. lithographed versus variable picture elements. refresh rate. brightness. density. color gamut. micro defects. aging. reliability. aliasing. artifacts. aspect ratio. overall display image area. display package size. power requirements. and cost. Some of these factors. not already discussed. will be expanded on.

There are two general categories of viewing environments: single viewer and multiple viewer. Traditionally. single viewer display sizes have been smaller (< 17 inches) and the applications have been more "task" oriented and interactive (e.g.. computer display). Multiple viewer displays have been larger (> 19 inches) and been more "entertainment" oriented (e.g.. TV) and passive.

Both of these traditions are changing and will continue to do so. In particular. a proliferation of single viewer entertainment displays (e.g.. personal TVs and games). and multiple viewer task displays (e.g.. electronic white boards and overhead projection panels) will be fueled by continuing advances in display technologies.

Viewing angle is an important factor in tuning displays to applications. The viewing angle is a function of display size and viewer distance. For a constant spatial resolution at the viewer's retina. overall display resolution needs to increase as viewing distance decreases.

As screen sizes increase and images get brighter. flicker becomes more of an issue in scanned displays. A 72 Hz scan rate can produce noticeable flicker with younger viewers in some situations. At a 50 or 60 Hz display scan rate. screens with high brightness that cover a wide field of view can produce objectionable levels of flicker.

It is important to separate flicker produced from scanning a display (commonly a flying-spot on a CRT). from other causes (e.g.. capture. conversion. interlace. signal processing artifacts).

Head mounted viewing devices (glasses or goggles) could make single user. low cost. high resolution displays practical. Additionally. a viewing device with dual displays (versus a mirror arrangement) would have the inherent ability to display stereoscopic images. This type of device could have both task and entertainment applications. would have a size and privacy advantage for portable applications (e.g.. portable computing. viewing proprietary videos on airplanes). and operate well in poor ambient lighting situations. It could also be the lowest cost way to deliver high resolution images to the early consumer market.

Future displays might also provide stereoscopic images without special viewing glasses and virtual holography (stereoscopic images with multiple viewing perspectives). Image architectures would need to pay particular attention to accommodating the latter.

As the market demands increasingly improved display capabilities. entirely new technologies and display structures may come into being. Some features which could find their way into future displays include: directly addressable image elements. layered structures with control over picture element persistence. variable spatial and temporal resolutions across the surface. on-display scene creation and manipulation. fixed eye position displays that map resolutions to match human retinas. eye tracking displays that tune resolutions across the surface to produce an optimal image to the viewer.

6.3.2 Practical Limits

A digital image architecture that is extensible and interoperable can be expected to improve in quality over the years. There are some practical limits beyond which human visual perception can be exceeded. making further gains in hardware capabilities nonproductive.

Although current sound technology parameters are close to or beyond human audible capabilities. there is a vast chasm of opportunities to be filled before we approach our optical limits. There are more mundane limitations as well.

For example: while one might imagine a 20 meter diagonal display with 0.1 mm pixel pitch, there are practical limits to both physical display size (how many home living/entertainment rooms could support such a large screen?) and human capabilities (one would need a magnifying glass or binoculars to appreciate such spatial resolution). On the other hand, close examination of images from the old masters might justify just such a display. And topological images might even require that this physically limited display be manipulated to bring in additional portions or viewpoints of the larger source image(s).

6.3.3 Future Receiver/Display Possibilities

During the next twenty years, some receiver/display options that should be considered include:

Wrist watch display

- single viewer, close viewing
- broad range of ambient light levels
- low pixel count
- low power
- wireless signal reception
- plug-in capsule with pre-recorded programming
- low to modest signal processing

Personal viewing device

- eyeglass "heads-up" display
- resolution matching the human retina
- stereoscopic image capable
- head tracking capable (signal bandwidth permitting)

Home entertainment display (classical HDTV)

- large size (two meters diagonal)
- viewing at moderate distances
- medium to high ambient light levels
- projection technology transitioning towards flat panel

Physician's work surface (simulates X-ray light wall)

- large size (two meters diagonal)
- close viewing/fine pitch
- optional touch screen input
- need to magnify small portions of multiple images
- composite of mostly still monochrome images
- large dynamic range of gray scale
- need to have shared display regions with other physicians
- can display other information

Engineer's white board

- large size (two meters diagonal)
- close viewing/fine pitch
- touch screen input
- composite of multiple image windows
- variable spatial and temporal resolution over screen sections

Drafting table

- large size (1.5 meters diagonal)
- close viewing/fine pitch
- pen/touch input
- monochrome acceptable
- low temporal resolution

Writer's work table

- moderate size (one meter diagonal)
- presents multiple "pages" of documents
- pen/touch input
- affords easy cut and paste (reducing paper drafts)
- text oriented, with some image capability
- close viewing/fine to moderate pitch
- monochrome acceptable
- low temporal resolution

Artist's canvas

- moderate size (one meter diagonal)
- close viewing/fine pitch
- excellent spatial features (resolution, contrast, color gamut)
- low temporal resolution (unless animation is required)
- specialized input/control/editing devices

Make-up mirror

- range of screen sizes and aspect ratios
- "through-the-screen" camera capability
- highly interactive image processing
- high temporal, spatial, color resolution
- color keying to cosmetics and lighting

Augmented imagery (ancillary to main viewing surface)

- manipulate ambient room lighting (lightning, color)
- manipulate moving room lights for peripheral vision effects
- physically animated room objects
- multiple viewing screens

6.4 Toward the AAAA (Anything, Anytime, Anyplace Appliance)

At the extreme in communications, imagine having a global archive, switching, communication, and receiving infrastructure in place that would let an individual access information in any media (Library of Congress, Smithsonian, encyclopedias, technical/professional journals, movies and TV programs from all film libraries, art from all major museums, company databases, home videos, personal medical records, new car facts with pictures and videos, weather maps, bank accounts, sports scores with selected replays, visual stock performance, restaurant menus complete with aromas and a view from your table, etc.) plus real-time access to any individual or group of people (voice, video, shared display, etc.)

7 Future Work and Other Issues

This Report is in essence a progress report on the requirements for establishing digital imaging standards that are interoperable, scalable and extensible. In many cases additional research and further discussions will be required to resolve the remaining issues. This section identifies some of those issues that need to be resolved. These provocative issues will benefit from discussion in the SMPTE engineering community at large. The SMPTE Standards Committee will determine engineering committee assignments to undertake a more detailed evaluation of these elements.

Gating Technologies —
A gating technology sets the pace for advancements in technology products and systems. For most of the history of television, the display was the fundamental gating technology. Only in recent years, has this role shifted to the transmission standard itself.

It should not be assumed that future architectures will be gated by display technologies over the long term. Other elements in the image path should be carefully evaluated as to their potential impact as gating technologies.

Strawman Proposals —
For a robust digital architecture standard to successfully cross industries, applications, and time, it is critical that thorough simulation and testing be performed across a range of applications.

At least three diverse strawman applications should be selected as test vehicles. Interoperability between these should be verified. Candidate applications include: broadcast television, multi-media computer workstations, medical X-ray, virtual reality, flight simulation, video phone, scientific visualization, and client-server networks.

Long Term Extensibility —
An accurate forecast of technologies and applications for the next fifty years is unlikely. However, a diligent evaluation of potentially relevant work underway in research laboratories throughout the world, and a careful study of anthropology and market demographics, should help in achieving long term extensibility.

Specifically, breakthroughs in image capture, display, communications, storage, and signal processing technologies could all have a profound effect on future image based applications.

Frame Rate Evaluation —
There is a need to gather existing data and perform experiments to determine desirable frame rates for a range of services including sports and other high temporal events. Bad numbers for conversion to or from existing standards should be explored so that they can be avoided.

Other Issues —
The questions which follow should be considered in the future work leading to a digital image architecture. They cover a wide range of topics and considerations. Some may ultimately prove to be important, others may become insignificant due to advances in technology.

As discussed in Section 2, the concept of scalability refers to the ability to extract higher and lower quality results from a common signal format. The concept of extensibility indicates the need to accommodate future enhancements in systems due to the rapid pace of technology.

Scalability and Extensibility require that many of the following areas have mechanisms for increasing and decreasing:

- Image resolution (7.1)
- Image temporal rate (7.2)
- Image layers, overlays, and windows (7.3)
- Compression quality level (7.4)
- Data rate, in relationship to image quality (7.5)
- Image luminance dynamic range (7.6)
- Image colorimetric range (7.7)
- Image, number of active channels (7.8)
- Audio quality (7.9)
- Audio, number of channels (7.10)

In each of the above areas, many issues arise in evaluating efficient mechanisms for increasing and decreasing.

7.1 Image Resolution

7.1.1 Are simple fractions a good guide for image resolution scaling and extending? If so, what should be the numerical basis?

The following issues support the use of simple fractions when scaling resolution:

a) The use of raster aligned overlay information (both real and virtual overlays), as is typical on computer screens.

b) The use of scalable compression techniques which scale in octaves.

c) The use of picture detail near or above the Nyquist bandwidth limit, for those applications where appropriate (typical in computer generated imagery).

d) The ability to use simpler and lower-cost transcoding in some situations, with less degradation (due to phase alignment).

e) The potential ability to get slightly improved transcoding (due to phase alignment) in high quality, higher cost transcoders.

7.1.2 Given that CCDs, active matrix liquid crystal displays and projectors, and other devices have inherently lithographed, and therefore very fixed, resolutions, should the numerical basis and specific (and optimally related) image sizes be definitively determined?

7.1.3 Further, in light of the many such fixed size, resolution, and raster format devices, would not square pixels be an important consideration? How many industries and applications require square pixels, and would be hurt if a non-square pixel format were chosen for advanced television? How severe would the degradations in quality and conversion cost be in such a case? Conversely, how many industries and applications not requiring square pixels would be hurt if a square pixel environment were imposed upon them?

7.1.4 If simple fraction resolution transformation guidelines are deemed worthwhile, should there be a numerical basis of certain base resolutions? Is the horizontal axis more critical for the basis due to digital design considerations? If so, are powers of two the optimal basis for horizontal resolutions, such as 512, 1024, 2048, etc.?

7.1.5 When transcoding resolution, what parameters are required to perform optimal transcoding? Is the bandlimiting associated with transcoding at ratios other than simple fractions an acceptable degradation? In what industries/applications is it acceptable, and not acceptable?

7.1.6 CCDs, active matrix liquid crystal displays and projectors, and computer generated images, and other image scanning, generating, and displaying devices can produce digital image values which are not bandwidth limited. Further, it is common for computer displays to use text, windows, and graphics which are aligned to the raster and which use maximum bandwidth signals such as white lines of pixels on black and black dots on white, etc. Given that such non-band-limited signals are common and useful in many industries, is the issue of requiring band limiting for transcoding, compression, or coding problematic? What industries would be significantly hindered if high definition systems required band limiting?

7.1.7 What useful increments of scaling might be best for a resolution hierarchy? Factors of two, being one optimally decodable resolution per octave? Two samples per resolution octave such as 3/4 and 1/2, or 3/2 and 2? Or is continuously variable resolution, and associated band limiting, a requirement in some industries/applications?

7.1.8 Should an image architecture emphasize the ability to apply more resolution to some screen areas than others? Or should constant image resolution and quality be mandated for all areas of the image? Is the answer different for different industries/applications? What problems might arise if such a signal format were considered for production? What issues arise within production switchers for such formats?

7.1.9 If some image areas are updated with different resolutions, or temporal rates, than others, should the universal header or descriptor contain this information and make it visible to all devices, or is it acceptable if such information is hidden within the data stream?

7.1.10 Are image region rectangular and square structures, such as that proposed as tiles, a useful construct in providing interoperability and flexibility in image update?

7.1.11 How likely are future image structures which are not *xy* raster based such as hexagonal or poisson distribution samplings? Is it possible to develop an image architecture which has mechanisms to accommodate such structures in the future? Can we anticipate the transcoding steps between a square-pixel *xy* raster and a uniform hexagonal or poisson distribution raster and thereby do our best to allow for such future possibilities?

7.1.12 How completely should image filtering and processing histories be specified in order to support subsequent image processing operations? A knowledge of the concatenation of all pre-filters may be desirable in complex image operations. How lengthy are such histories likely to become?

7.1.13 How do flying-spot devices such as CRTs and cathode-ray cameras relate to fixed raster "lithographed" devices such as CCDs and active matrix liquid crystal displays and projectors? How can the high definition image architecture accommodate both types of image sources and displays without substantial quality loss? Can both types of image data be processed in the same transforming devices using the same parameters, or are different processing steps required for the two different types of image data?

7.1.14 Is there a representation appropriate in which an idealized pixel can be generated through signal processing? The purpose of such an idealized pixel would be to be used as input to resolution scaling (also known as resolution transcoding). Would the digital signal processing required to create such an idealized pixel result in unacceptable artifacts due to the footprint of the convolutional processing kernel?

7.1.15 In traditional television, there was no possibility of color dot triad alignment with scanlines or pixels. However, with the advent of lithographed displays, such as active matrix flat panel displays, the relationship between the color triad and the pixel becomes exact. Is there an optimal organization of color area portions in the context of a digital image architecture? Should the colors be overlayed onto a common area through the use of lenslets, fibers, or other techniques? Should the pixels be adjusted so that they even overlap through such optical techniques in order to reduce blocky appearance?

7.1.16 If the color regions representing a pixel must remain spatially distinct, is there a particular arrangement which is optimal? If so, should the precise positions of the color sub-pixel areas be taken into account in the digital image architecture and in the representation, capture, and processing of the digital image signal? What effect do the gaps between color regions have? How do lenticular or Fresnel screens affect color?

7.1.17 If the color regions representing a pixel must remain spatially distinct, could more than one triad be placed within one logical pixel? Is there an optimal number of such sub-triads, such as perhaps four? Is there an optimal arrangement of such sub-triads? If so, should this arrangement be taken into account in the digital image architecture and in the representation, capture, and processing of the digital image signal?

7.1.18 If a color space were to support more than three primary colors, would there be benefit to using four or more primaries in some appropriate configuration as a standardized pixel shape?

7.1.19 On a CRT, the spot shape is usually a round or ellipsoidal gaussian which flies horizontally. On a flat panel display, the spot is usually a square shape which is stationary. There may also be dead-zones between pixels. What signal properties should be adjusted to take these issues into account?

7.1.20 Could some of these issues be handled by the use of an idealized or standardized pixel representation, with defined transformations at the receiving device appropriate to its particular pixel configuration? If so, how should the digital image architecture specify this?

7.1.21 What should be done concerning similar issues in CCD image sensors?

7.1.22 What is the impact of these various issues on the Kell factor?

7.1.23 Would other pixel shapes such as triangle, hexagonal, and diamond have advantages for future image capture and display technologies?

7.1.24 The 1.333:1 (4 x 3) aspect ratio is widely used in television and computers. In the motion picture industry, 1.37:1, 1.66:1, 1.85:1, and 2.35:1 are all commonly used. The 8.5" x 11" page has an aspect ratio of about 0.77:1. The European page size approaches 0.71:1. Computer display memories are most simply organized with aspect ratios which are simple fractions such as 1:1, 3:2, and 2:1. Medical radiology, still photography, newspapers, magazines, books, and other images have a number of commonly used aspect ratios. How do we achieve support of all of the aspect ratios in common use?

7.1.25 Can the header/descriptor be used to indicate the aspect ratio and resolution of an image, so that the displaying device can do its own version of letter boxing (unused areas) or overscan (discarded areas)? For those systems which have compressed frame groupings, how could the edges of the letterbox be protected from the moving blocks?

7.1.26 What would be the most widely used mappings between common aspect ratios and anticipated common screen sizes?

7.1.27 If the 16:9 aspect ratio becomes popular, what mappings are likely for European (metric) and American (English) sized paper pages, wide screen movies at 2.35:1, television and film at 1.333:1 (4 x 3), movies at 1.85:1, and other widely used image formats?

7.1.28 Can the digital image architecture support not only a wide variety of aspect ratios in the material being displayed, but also a wide variety of aspect ratios at the receiving display itself?

7.1.29 If the digital image architecture supports multiple aspect ratios, with interoperability between displays at such various aspect ratios, what are the key technical issues? Should the horizontal resolution of all aspect ratios be held at simple fractional relationships, while allowing the vertical resolution (with square pixels) to vary in fine increments to fit the exact aspect ratio desired at the display?

7.1.30 The diagonals of the camera apertures of common film formats have dimensions as follows:

Format	Diagonal
35mm Full Aperture	31.14 mm
35mm Academy	27.16 mm
35mm Still	43.27 mm
65mm	57.30 mm
Professional Roll Still	101.1 mm

There are many high quality lenses in existence for these formats. Would it be useful to keep these dimensions in mind when developing lithographed sensors such as CCD arrays?

7.2 Image Temporal Rate

7.2.1 Should a single temporal rate be emphasized? Or is there a need to support multiple temporal rates for different industries/applications, or within one application?

7.2.2 Are there sufficient mechanisms available in temporal properties of high definition systems to handle the issue of computer CRT displays requiring refresh rates higher than 70 Hz?

7.2.3 Is there a mechanism for reliably and consistently transforming high definition television imagery to 24 fps film for theatrical release?

7.2.4 Should there be a family of temporal rates which are related by a simple fraction rule? If so, what should be the numerical basis of such rates?

7.2.5 When temporally transcoding, what temporal beat frequencies are visually acceptable? Does the 12 Hz beat frequency of the 3-2 pull down, and its wide use and seeming acceptance, indicate that 12 Hz or higher is an acceptable beat frequency, or are there frame patterns in which higher beat frequencies are required for acceptable viewing?

7.2.6 What sort of synchronization mechanisms are optimal for digital systems, given that inherent digital system flexibility need not require every device to be locked to a common master very-high-frequency oscillator (near 100 MHz)?

7.2.7 CCDs, active matrix liquid crystal displays and projectors, and other devices, utilize a portion of horizontal or vertical retrace intervals to transfer to/from frame buffers. Future systems may not require these intervals. How will this affect the need to dedicate signal time to these intervals?

7.2.8 Given that CCDs, active matrix liquid crystal displays and projectors, and other devices, have no inherent flicker or update rate requirements, should temporal rate flexibility be part of the high definition architecture?

7.2.9 Should an image architecture emphasize the ability to use a higher update rate for some screen areas than others? Or should constant image rate be mandated for all areas of the image? Is the answer different for different industries/applications?

7.2.10 It is common to use a 50% temporal duty sampling cycle (180 degree shutter in film cameras to allow film pull-down), which provides a balance between motion blur and sharpness. Is not this temporal undersampling certain to introduce temporal aliasing during temporal rate transcoding? Is not such aliasing certain to appear as artifacts which occur at the temporal beat frequency rate (e.g., a 50 Hz to 60 Hz transcoding would have a 10 Hz beat frequency)?

7.2.11 Some CCD sensors used in cameras see the entire frame area during the exposure time. This is similar to film exposure. Some tube cameras scan the image top to bottom, whether progressively or interlaced. What are the temporal processing, displaying, and viewing effects caused by mixing devices which integrate the entire image versus those that scan the image from top to bottom? What temporal issues arise due to the fact that the top of the image may be seen or displayed nearly a frame time before the bottom of the image, and half a frame time before the center of the image?

7.2.12 As just mentioned, both displays and sensors exist which scan from top to bottom or which integrate the entire image for the frame time. What architectural issues should be examined in attempting to take this issue into account? What issues are involved in converting a scanned image for area display or in converting an area sensed image for scanned display? How do these issues affect film scanning such as in a telecine? Does the wipe time involved in the physical film camera shutter have an affect? How do these issues affect film recording from an electronically captured moving image?

7.2.13 How do these scanning pattern issues affect temporal transcoding, finding motion vectors for compression or standards conversion, effects processing, or other image processing operations? What issues arise when compositing or mixing multiple image sources captured with different scanning patterns?

7.2.14 Standards converters which convert between 525/60 and 625/50 are available at various levels of cost/performance, utilizing a number of techniques. At the highest levels of performance, motion estimation may be employed to interpolate frames. Undersampling due to interlace may have an impact of this process. What problems may thus arise for architectures that rely on temporal transcoding or standards conversion?

7.2.15 When displaying multiple windows of moving images on a screen, as in a future video teleconference, how can buffering be minimized for each picture stream in order to achieve display synchronization? What options are available for local, regional, and global synchronization, both loose and near exact?

7.2.16 Is there a benefit to selecting a particular master oscillation rate, from which pixel clocks in the scalable system are derived? If so, what candidate rates might offer advantages?

7.2.17 Some applications, such as teleconferencing, interactive flight simulation, or virtual reality, require low latency. Other applications, such as broadcast television, can have substantial latency without much problem.

What digital image architecture mechanisms are needed to provide for those applications which require low latency?

7.2.18 Compression algorithm design is significantly affected by a low latency requirement. What latency is implied by any candidate digital advanced television system? What affect would such inherent latency have on usefulness for those applications requiring low latency?

7.2.19 Digital network design is affected by needs for real-time bandwidth as well as latency requirements. How does the need for low latency combined with high real-time bandwidth in these industries affect digital interactive network design?

7.3 Image Layers, Overlays, and Windows

7.3.1 In some workstations, one to four bits (usually two or four) are used for overlay planes, which are added on top of the underlying image (which is usually full color). However, many popular systems, like the Macintosh II®, the DECstation 5000®, and others, do not use overlay planes. It is possible for X-Windows® and other window management systems to support overlay planes as well as managing full color windows in the main bit planes? Should overlay planes be part of digital advanced television architectures?

7.3.2 If overlay planes are used, should these overlay planes be implemented in hardware, or as a virtual mechanism in software, or can both be accommodated?

7.3.3 How many bits of real or virtual overlay plane should be mandated or recommended, if overlay planes are mandated or recommended?

7.3.4 The common practice of bandwidth limiting moving images suggests the possibility of using overlay planes to contain non-band-limited imagery, with the band-limited moving images using the main bit planes underneath. Overlay planes could easily contain the usually non-band-limited window borders, text, stipple patterns, graphics, etc., which characterize computer screens. This architecture would require all receiving devices to support either real or virtual overlay planes. Is such an architecture appropriate?

7.3.5 Is it necessary for interoperability across industries and applications to allow for the possibility of non-band-limited picture data co-existing with band-limited imagery from cameras or other (possibly synthetic) sources?

7.3.6 Do appropriate digital image compression algorithms exist which can pass non-band-limited picture data such as that used typically on computer screens? If such a compression technique exists, would this allow such non-band-limited picture data to be transmitted together with the moving picture stream? What are the properties of such a compression algorithm, if one exists?

7.3.7 Can the data areas available in some digital advanced television proposals be used to convey encoded data for use with real or virtual overlay planes? Would Unix X-Windows®, Display Postscript®, Apple Macintosh Toolbox®, Microsoft Windows®, fax, or other forms of encoded graphic and text data such as run-length codes be conveyable in this manner? Are there one or more such techniques which might be appropriate to support for digital advanced television?

7.3.8 Is the proprietary nature of many of these formats a barrier to interoperability, or are there potential solutions to provide universal access?

7.3.9 Are open standards such as IGES for vector and graphics images, CGM for raster images, or Open Document Architecture (ODA) for compound documents worth considering in light of desire for universal access?

7.3.10 Should one or a small number of such formats be supported universally? By standardizing on only one such format, all receiving devices would only need to support that single format. If no such single standard is chosen, then each receiving device desiring to display computer-type window or graphics displays might need to support many or all of the formats in common use. Is there a way to encourage adoption of one or a small number of text and graphics protocol standards to be universally supported?

7.3.11 Should a digital image architecture require that text and graphic data, typical of computer displays, be able to be passed to the display by either the use of real or virtual overlay planes or appropriate compression algorithms capable of passing this information? Should the digital image architecture insist on at least one of these two ways of passing computer display information?

7.3.12 As an alternative to screen-resolution-specific graphics, should all graphics be specified with much higher precision than the display? Such might likely use outline fonts and graphic commands which can do proportional blending of text with appropriate filters, and which allow image detail to placed between pixels or lines. Would such non-raster-aligned text and graphics, with appropriate filtering, be acceptably legible and clear compared to raster-aligned and non-band-limited text and graphics as is typical of current computer screens?

7.3.13 If presentation of non-band-limited image information, as is typical of computer displays, is a requirement, then should multiple screen resolutions be supported for resolution scalability? If so, should the simple fraction guideline be used for the relationships of

screen resolutions due to the need to preserve legibility and clarity of the non-bandwidth-limited image data?

7.3.14 Should the capability for selecting among a variety of overlay-planes for display be part of digital advanced television architectures? Could such a selection be useful for closed captions, sign-language inserts, foreign language subtitles, television program guides, sports statistics, or other picture augmentation information?

7.3.15 How many such simultaneous overlay planes could or should be supported?

7.3.16 Should more general compositing functions other than overlay be supported at the receiving device as a part of digital advanced television architectures? In particular, should alpha blending be supported? Alpha blending uses proportional blending at edges of the overlaying area. This technique is also known as proportional matte edge compositing.

7.3.17 If such compositing should be supported, should there be limits or scene-specification guidelines for the amount of area involved in mattes and proportional edges each frame, or each second?

7.3.18 If such compositing should be supported, how many layers of composite should be allowed? If the activation of the composited foreground overlays is supported, it is likely to be controlled and specified by the user at the receiving device. If foreground overlays are transmitted in moderate time intervals such as a second or more, then the number of overlays allowed will affect the amount of buffering required at the receiving device. Should such issues be a part of an image architecture for digital advanced television?

7.3.19 Is the concept of tiles and plates useful by providing compact data representation for locating pixels used in proportional alpha blending? Tiles and plates are screen area subdivisions allowing specification of screen locations. The concept of plates is similar to city blocks and the concept of tiles is similar to house lots, with pixels being similar to locations on a grid laid over the house lot. This organization makes addressing a given pixel, tile, or plate location simpler. It also allows all locations on a given street to be easily located due to the proximity of each house to the next. Is this method of area subdivision useful? Are there other methods for subdividing an image which are appropriate or useful in digital advanced image architectures?

7.3.20 Should the use of windows on the display of the receiving device be anticipated as part of the architecture? What potential affect would there be on the architecture by anticipating the use of two or more windows within the display, each independently controlled by each user at each display? Is the minimal functionality of picture-in-picture appropriate, or is the elaborate window sizing, positioning, and overlaying capability of typical computer window systems appropriate, or somewhere in between?

7.3.21 If windows are anticipated, how would dynamic resizing take into account the possible need for simple-fractional guidelines in resolution scaling? Should there be notches in the window sizes at simple fraction points to allow clearer and more legible text and graphics?

7.3.22 If the local computer wishes to use the overlay planes, how would such use interact with remote control of these overlay planes? Would the local system have priority? How would such priority be controlled?

7.3.23 Is it not likely that locally simulated synthetic computer-generated images will often be overlayed onto the digital advanced television image stream? What filter representations are appropriate for matching the pixel representations of the simulated and real or television imagery? Should idealized pixels be used? What should be the assumptions concerning flying-spot versus lithographed raster pixel representations in this context? What are the optimal computer anti-aliasing filters for such composited images?

7.4 Compression Quality Level

7.4.1 What is the proper tradeoff between compression quality, data rate, image resolution, image dynamic range, image color fidelity, and image temporal rate? How should these tradeoffs be resolved for different industries and applications (e.g., medicine, scientific visualization, production, videoconferencing, transmission)?

7.4.2 Can a family of compression quality levels be developed which allow spatial and temporal resolution scalability with a layered coding technique? Can such a compression technique compete successfully with single point solution or single resolution/rate heavily optimized systems? Would such a technique offer benefits to interoperability between industries/applications? Are the benefits of scalability and extensibility sufficient to justify the effort to develop a layered compression system of sufficient quality? How much is gained by compromising the orthogonality required for scalable layered systems via the use of non-linear terms? Are such non-linear data interactions likely to hinder other interoperability needs, in addition to the desire for scalability and extensibility?

7.4.3 How serious is the problem of motion vector interaction between motion vectors used in standards conversion and motion vectors used in motion-compensated compression? Would an image which has been transcoded in either resolution, temporal rate, or both, interact acceptably with a subsequent digital compression algorithm, or would the compression motion vectors

show severe errors and aliasing beat frequencies? If the image is further transcoded after subsequent decompression, would the errors compound further?

7.4.4 What sub-pixel resolution is required for motion vectors used in motion-compensated compression?

7.4.5 What is the effect of concatenated compress/decompress cycles within one algorithm as used when exchanging images multiple times between different industries/applications? What is the effect of concatenated compress/decompress cycles between different compression algorithms? Are the preliminary studies which indicate that this might result in severe degradation correct?

7.4.6 It is common practice in some compression schemes, such as MPEG, to use frame groupings. How can live switching be performed if frame groupings are not aligned? How can misalignment of key intermediate anchor frames be prevented when performing multiple compress and decompress cycles?

7.4.7 Do these issues argue for commonality and compatibility of compression algorithms, and a minimization or elimination of temporal and spatial transcoding in processing images? If transcoding is applied, are simple fraction-based temporal and spatial transcodings less prone to degradation than arbitrary fraction transcodings?

7.4.8 Can a high resolution system architecture accommodate the rapid algorithmic and digital hardware advances in the state of the art which appear to be inevitable each year? If an optimal algorithm for this year is selected, what is the likelihood that this algorithm would be obsolete in five or ten years due to improvements in hardware or algorithmic techniques? If compression algorithms are likely to become obsolete every five years, what high definition system architecture principles can be developed to allow radical algorithmic and hardware improvements as appropriate? Is the header/descriptor sufficient, or are other principles required in conjunction with the basic compression algorithm design to allow extensibility or easy replacement/upgrade? Can new algorithms coexist with old algorithms while providing efficient bandwidth/spectrum usage, in light of the fact that new algorithms may be many times more efficient than old ones? Can old algorithms be continued in use when their inefficiency approaches factors of four or eight below optimal with respect to the newest algorithms and hardware? How can extensibility be accommodated, as will certainly be required, while maintaining service to older devices which require inefficient digital signal architectures for a given point in time?

7.4.9 How likely will the current DCT and sub-band systems advance as future optimal algorithms in five to ten years? Are other algorithms such as fractal, wavelet, vector quantization or other large codebook algorithms likely to be more efficient in some future hardware capability level? Is it likely that future compression algorithms may be as yet unanticipated? What steps can be taken to prepare for such major shifts in compression techniques, should they occur?

7.4.10 Is it likely that decompression chips in receiving devices could be programmable? If so, could updates to compression algorithms be downloaded using header/descriptor support, or by other software distribution methods? Would such updates be useful? Would it be useful to place the decompression module on a standardized card, so that the chip itself could be replaced as technology advances?

7.4.11 Are there proposed signal formats based upon very rapid partial frames which can support multiple receiver display rates from a single signal without degradation of any rate? Do such formats also provide for minimum buffering of multiple asynchronous sources being presented on the same display?

7.4.12 Some applications, such as the colorization of movies, create data which defines the objects and their boundaries and motion for every frame. Can such data be useful in a system architecture for compression or other uses?

7.4.13 Is it useful to gather macro information about a scene by encoding data such as camera position, motion, and orientation? Could tripod head encoders be useful for this purpose? Are there navigational tools which could be adapted for this purpose? Would such global information be of sufficient value in compression and other uses to warrant the capturing of this information?

7.5 Data Rate in Relationship to Image Quality

7.5.1 Can digital image compression algorithms be layered in resolution and temporal rate, while at the same time being usefully layered in data rate?

7.5.2 Is a data rate hierarchy possible in this context?

7.5.3 There will always be a variety of bus rates, memory bandwidths, disk transfer rates, and channel rates. What are the useful rates of a data rate hierarchy, if such a hierarchy is possible?

7.5.4 Is orthogonality of temporal and spatial resolution via a layered hierarchical compression technique possible? Is data rate orthogonality, integrated with temporal and spatial resolution, possible?

7.5.5 Can other useful augmentations be layered onto the data rate such as extra camera views, stereoscopic imagery, z-value depth information, blending coeffi-

cients for compositing, additional dynamic range or improved colorimetry?

7.5.6 Can high quality still frames be acceptably interleaved into the data stream concurrent with the moving image data stream?

7.5.7 Can alternate aspect ratios be provided simultaneously by an appropriately layered data stream construction?

7.5.8 Can extra image channels such as closed-caption sign language display windows, previews of future shows, and others, be acceptably layered into the data stream?

7.5.9 Can three dimensional image construction information be provided for those receiving devices capable of creating three dimensional computer generated images? Could such images, by computer synthesis at the display, substitute for tiny details which are not adequately captured with the camera resolution such as red-orange golf balls (usually red and blue have lower resolution than green due to chroma-sub-sampling in the Y,Cr,Cb technique)? Could computer graphics create new interactive games or other locally interactive education or training in this way? Will future display devices be likely to have the capability to generate some amount of screen area containing three dimensional computer generated synthetic images and composite them correctly into the two-dimensional high definition background image situation?

7.5.10 Transport errors are highly dependant on the transport channel. What sort of error protection/correction should accompany advanced television digitally compressed data, in light of such data's extreme sensitivity to errors? Are the mechanisms being developed in the transport header portion of the universal header/descriptor sufficient, or should all digitally compressed picture data contain inherent protection and correction protocols?

7.5.11 How should encryption be supported? Is public-key encryption appropriate and sufficient? What levels of encryption are required for various transport media and various uses?

7.5.13 Packet-retry type networks, such as the current Internet, or Ethernet® with TCP/IP cannot guarantee delivery of data for a real-time stream since a packet of data may be "bumped" and must be resent. Real-time streams require that data be both intact as well as "on-time," invalidating packet-retry protocols which would provide resent data subsequent to the required time. It is therefore likely that the basic data network infrastructure will need to significantly change in order to support real-time imagery and audio streams in shared channels, unless switched point-to-point services are used due to insufficient shared-channel infrastructure technology. In light of such significant changes, is it possible to anticipate the future networking protocols and techniques so that the high definition image architecture can be developed to be compatible? Are packet prioritization and priority-based graceful degradation likely to be key techniques in such future networks? If so, how much priority information and priority verification and authorization is needed? Also, how many levels of priority, and what priority schemes might be required, to optimize quality for all users of a shared channel as well as priority packet routing performance? How much guaranteed bandwidth is required by each type of user, and can such data bandwidth be guaranteed? Will such potential guarantee requirements need some amount of data bandwidth reservation? Is it possible to design appropriate shared networks with a hybrid of reservation and non-reservation, applied with both reservation and non-reservation portions of each connection's data stream? Is such a technique a reasonable match between prioritized compressed advanced television and shared data networks, such that a certain amount of real-time bandwidth is guaranteed, but an additional portion is not reserved but is usually provided?

7.6 Image Luminance Dynamic Range

7.6.1 As cameras and displays continue to increase their luminance dynamic range beyond the 100:1 which is now common, would it not be desirable to be able to support such extended range? Are 1000:1 luminance dynamic range cameras and display devices likely in the next five to ten years?

7.6.2 Can a given advanced television digital compression algorithm be augmented to allow more bits for luminance, or would the requirement for more accuracy defeat the ability of the algorithm to provide the required compression ratio?

7.6.3 Can luminance transfer functions such as those used in CCIR Rec. 709 and SMPTE 240M be augmented to provide extended black and white range?

7.7 Image Colorimetric Range

7.7.1 Do there exist cases in which red or blue details on dark backgrounds, or yellow or cyan details on light backgrounds, would require higher resolution in color for both broadcasting and other industries/applications?

7.7.2 What are the benefits of an RGB digital representation for interoperability across industries and applications versus the Y,Cr,Cb (also called Y,Pr,Pb and YUV) color difference representation commonly used in television?

7.7.3 What are the tradeoffs for using wider gamuts, including gamuts beyond the real spectrum, in covering

the real colors? What industries/applications, such as perhaps museums, colleges, printing, photography, and motion pictures, may require accurate color reproduction over a color gamut which is wider than is commonly proposed for high definition television systems?

7.7.4 What other color representations, such as HSV, CIE x,y or CIE u',v' are useful?

7.7.5 How much precision loss accompanies a given high definition color transformation to and from these other representations?

7.7.6 Is there benefit from using a color space which supports color sensors and displays which use more than three color primaries?

7.7.7 When are linear color representations needed for computations? What linear color representations are appropriate in a device-independent context?

7.7.8 What color representations offer device independence?

7.7.9 Can luminance or other brightness measure be represented such that it is orthogonal to color representation? Can such a representation offer color invariance under exposure or illumination level adjustment?

7.7.10 What color representation is most useful for adjusting a wider gamut of color for a narrower gamut display? What are the tradeoffs between clipping to the narrower gamut, and a softer adjustment, similar to highlight compression in the *S*-curve?

7.7.11 What are the perceptual uniformity properties of various useful color spaces?

7.7.12 Are Hue, Saturation and Value representations useful in these contexts?

7.7.13 How can digital numeric representation efficiency be optimized, while still allowing the possibility of wide gamut colorimetry in addition to efficient support of narrower gamuts.

7.7.14 What device independent color spaces are most efficient for compression?

7.7.15 Luminance, which matches human color sensitivity, is an appropriate representation of brightness near the display. Other representations may be required in the studio where processing is required. For example, blue screen compositing involving transparency requires as much detail in blue as in green. How should the division be made between the use of luminance versus an equal representation of red, green and blue?

7.8 Image, Number of Active Channels

7.8.1 As efficiency of compression improves, if ways are found to allow new compression techniques to replace old ones, then new data bandwidth will be made available with each such upgrade. This bandwidth can carry more active image channels. Such channels can be of quality equivalent to the original if a factor of two is gained in the upgrade. This would make possible the two required image streams for stereoscopic images. Alternatively, or additionally, a number of lower-quality channels could be added with alternate views of the scene from different cameras (cockpit camera in a car race or helmet camera in football, closeups, long shots, etc). Also, entirely different programming is possible on the new data. Can an architecture be developed which will allow this evolution?

7.8.2 Is the header/descriptor mechanism likely to be a major element in providing such augmented capabilities, or do the issues extend into the nature of upgradable compression algorithms?

7.9 Audio Quality

7.9.1 Musicam (used in MPEG) and other compressed audio systems are also likely to improve. Similar issues of upgradability and backward compatibility must be considered when evaluating the system architecture. Can the image architecture for this upgrade path, if one is found, be applied as effectively to audio?

7.9.2 As compression techniques improve, a given level of quality can be maintained while reducing data bandwidth requirements. What are the best uses for newly freed bandwidth? Will older devices be able to decode new algorithms if they are made to be somewhat programmable from the beginning? Programmable devices will allow algorithmic improvements which can utilize a given device. However, it is likely that improvements in algorithms will have to be accompanied by new hardware, thus making the upgrade path and backward compatibility path difficult. Are there ways to improve this situation in the system architecture?

7.10 Audio, Number of Channels

7.10.1 Additional audio channels can be added in expanding bandwidth due to more efficient algorithms or due to increased bandwidth or higher reliability channels.

7.10.2 Likely uses for additional channels are six-channel surround sound, and multiple languages on separate tracks.

8 Annex

8.1 Glossary of Terms

Alias – a form of image distortion associated with spatial and temporal filtering. A common form of aliasing is a stairstepped appearance along diagonal and curved lines. See Scaling:Interpolation.

Compression – the process of removing redundancies in a digital data stream to reduce the amount of data that must be stored or transmitted. The following terms are often used in describing image compression systems:

Lossless Compression – techniques for data reduction in which the original information can be recovered exactly as it existed prior to encoding.

Lossy Compression – techniques for data reduction in which some information may be lost in the process of encoding or decoding the data. In image compression, an effort is made to preserve as much of the visually-significant data as is possible, sacrificing, when necessary, only data unlikely to be perceived by the average viewer.

Scalable Coding – the ability to encode a visual sequence so as to enable the decoding of the digital data stream at various spatial and/or temporal resolutions. Scalable compression techniques typically filter the image into separate bands of spatial and/or temporal data. Appropriate data reduction techniques are then applied to each band to match the response characteristics of human vision.

Fixed Data Rate Compression – techniques designed to produce a data stream with a constant data rate. Such techniques may vary the quality of quantization to match the allocated bandwidth.

Variable Data Rate Compression – techniques designed to produce a data stream with a variable data rate. Such techniques typically maintain a constant level of quantization producing a variable data rate based on the spatial and temporal energy content of the images being encoded.

Conditional Replenishment – a technique whereby various portions of the image are updated at differing rates. Interframe coding techniques utilize this concept as they eliminate redundancies between frames, only storing or transmitting image data related to the changes between frames. Conditional replenishment occurs as image sequences are reconstructed in dual-ported frame memory; current display designs typically scan the frame memory at a constant rate to refresh the display. The following terms are related to the concept of conditional replenishment:

Addressable Display – a display designed to allow each pixel element to be controlled independently, allowing for conditional replenishment.

Conditional Acquisition – an image acquisition system that processes the image data and outputs information about changes as they occur, up to the maximum temporal sampling rate of the image sensor. Parallel outputs from separate image regions (tiles) may permit higher sampling rates.

Incremental Update – the practice of sending information only about changes to an image that occur between temporal samples; computer graphics and animation systems typically employ this technique.

Multirate Encoding – the digital encoding of multiple image components that are presented to the encoder at different temporal rates. For example, a high spatial/low temporal resolution component may be acquired at 24 fps while a low spatial/high temporal resolution component may be acquired at 72 fps.

Latency – the delay (latent period) between the occurrence of an event and its display by an electronic imaging system. Each of the following factors can contribute to the total latency of an imaging system; in some cases image distortions may be related to the latency of the system:

Encoding/Decoding – interframe coding techniques utilized for digital video compression require the processing of multiple frames to eliminate redundancies in the static areas of the image and to calculate motion vectors for the motion components of the image. Depending on the sophistication of the coding, interframe compression can introduce delays from two to several hundred frames.

Storage – used here to describe the access speed of a digital image storage system. Fast access speeds reduce the time required to retrieve specific data. CD-ROM is said to have high latency due to its slow access speed, while hard disks have lower latency due to their faster access speeds.

Synchronization/processing – the use of frame synchronizers, timebase correction, and digital processors for visual effects introduce frame delays that accumulate as the signal passes through a video production system.

Pixel – the smallest picture element of an image (one sample of each color component). A digital display is typically specified in terms of pixels and color depth: the number of digital bits stored per pixel. A picture element is also called a *pel* in the field of image processing.

Quantization Levels – the predetermined levels at which an analog signal can be sampled as determined by the resolution of the analog to digital converter (in bits per sample) or the number of bits stored for the sampled signal. See: Sampling.

Resolution – The capability of an optical system, or other imaging system, of making clear and distinguishable the separate parts or components of an object. With respect to the relationship of the human visual system to an imaging system display, several factors must be taken into consideration:

Spatial Resolution – the ability of the display to reproduce adequate detail to allow the visual system to distinguish the separate parts or components of an object.

Temporal Resolution – the ability of the display to reproduce adequate detail to allow the visual system to distinguish the separate parts or components of an object that is moving through the display.

Perceived Resolution – from the observer's viewpoint, the apparent resolution of a display. This concept is based on the ability of the viewer to resolve all image detail presented by the display. At the ideal viewing distance, perceived and actual spatial resolution are equal; at greater viewing distances the perceived resolution is higher than the actual spatial resolution of the display.

Sampling – the first step in the process of converting an analog signal into a digital representation. This is accomplished by measuring the value of the analog signal at regular intervals called samples. These values are then encoded to provide a digital representation of the analog signal. Image samples are usually called pixels. See: Quantization Levels.

Scaling (spatial) – alteration of the spatial resolution of an acquired image to decrease or increase the number of pixels used to represent the image. Any of the following techniques may be used for image scaling, resulting in the addition of image artifacts as indicated:

Interpolation – the process of averaging pixel information when scaling an image. When reducing the size of an image, pixels are averaged to create a single new pixel that replaces two or more adjacent pixels; when an image is scaled up in size, additional pixels are created by averaging the values of adjacent pixels. Interpolation generally causes an apparent softening of the image when it is increased in size, because the averaging process does not create any new information.

Pixel Replication – a process used to display an image at a larger size by repeating pixels along a horizontal line and/or repeating lines to increase the vertical size. For example a 320 by 240 pixel image can be displayed at 640 by 480 size by duplicating each pixel along a line and then repeating the line; the resulting image will contain blocks of four pixels with the same value.

Resampling – the process of converting images between the spatial resolutions utilized by different imaging systems. The process may include interpolation to correct for differences in pixel geometry or to scale the image. Resampling is frequently used to change the density of pixels, typically measured in dots per inch (DPI), when preparing images for printing using halftones and color separations.

Sub-Sampling – bandwidth reduction techniques that reduce the amount of digital data used to represent an image. The following techniques are commonly utilized:

Chroma Sub-Sampling – the reduction of color resolution by reducing the bandwidth of color difference signals as practiced in composite video transmission and recording systems or by eliminating some color difference pixel information in digital processing systems.

Decimation (pixel sub sampling) – the process of discarding complete samples. The resulting image is reduced in size and may suffer from aliasing.

Scaling (temporal) – alteration of the temporal resolution of a visual sequence, to decrease (typical case), or increase the amount of data used to represent the visual sequence. This process may include one or more of the following techniques:

Interpolation – the process of adding or deleting temporal samples to a visual sequence by averaging adjacent temporal samples. Results are poor when there is rapid motion.

Motion compensation – the process of adding or deleting temporal samples to a visual sequence by predicting where a moving object should appear in the resulting new temporal sample.

Replication – the repetition of temporal data to increase the temporal display rate of a visual sequence. Typically used in computer systems to provide a higher screen refresh rate than the temporal rate of the visual sequence being displayed. Also used in film to video conversions to change from the 24 fps temporal rate of film to the 30 fps temporal rate of NTSC composite video (3-2 pull down).

Speed Change – playing information acquired at one temporal rate at a different temporal rate. For example, 24 fps film is played at 25 fps when it is converted to PAL composite video (a 4% speed change).

Sub-sampling – dropping entire temporal samples (e.g., video frames) to reduce the data rate.

8.2 Temporal Rate Analysis

This section includes detailed descriptions of some techniques for temporal rate conversions that might be utilized in the process of translating existing film or television pictures for display at higher refresh rates.

8.2.1 59.94 and 60.0 Hz versus 72 and 75 Hz

Computers which use CRT displays (as opposed to flat panel displays) currently comprise nearly all of the market except for portable laptop computers. Substantial experience with high resolution image presentation, with wide field of view and high brightness, has increasingly led to higher refresh rates such as 72 and 75 Hz.

The use of 72 Hz is very naturally compatible with 24 fps film, being exactly three times. The rate of 75 Hz is compatible with the common European practice of transferring motion picture film to 50 Hz video by running it at 25 fps or 4% fast. A display which has a sync tolerance range of 4% could adapt to both 72 Hz and 75 Hz picture rates. A 4% tolerance does not add undue cost to a display. A CRT or other display which can present both 72 and 75 Hz progressively scanned images can be a key architectural element in a digital image architecture. Such a display can be used on computers as well as being able to be used for high quality presentation of motion picture film, and for 50 Hz material.

8.2.2 Motion Prediction

Very expensive receivers can attempt to perform motion-predictive frame rate conversion in the receiver. The motion vectors which are used in the motion compensation portion of the digital image compression process could be helpful in this regard. Compression only requires statistically beneficial correlation between the motion vectors and image motion, while rate conversion requires an accurate motion analysis in order to be artifact free.

8.2.3 Temporal Undersampling

It is common practice to temporally undersample. This leads to reverse wagon wheels and many other well known temporal aliasing artifacts. In motion picture film production, temporal undersampling takes the form of a near 50% duty cycle camera. Shutter angles near 180 degrees are typical for most production, where approximately half the time is used to expose each frame, with the other half being used to pull the film down to the next frame.

In video, the temporal light sampling time for each pixel is sometimes adjustable. Video temporal light-capturing duty cycles will vary from about 30% to about 100%. Short duty cycles or adjustable duty cycles are only available on some cameras, since light sensitivity is usually reduced. However, even some home video cameras have a persistence type control to allow image capture duty cycle adjustment between sharp frames and smooth motion.

Some experts feel that a 50% exposure duty cycle for each frame is the most common choice since it provides a balance between image sharpness and smooth motion. It is likely that some photography of moving objects may desire a short shutter time in order to emphasize sharpness, and to distinguish each image on each frame separately. Other uses such as "go motion," used successfully in many Lucasfilm productions, may favor smooth motion by choosing a 100% duty cycle. The 100% duty cycle is achieved in this example by a repeat motion non-real-time computer controlled camera.

In computer graphics, when motion blur is simulated, a duty cycle is usually specified. Some standard software rendering packages, such as Pixar's Renderman®, offer a value from 0 to 1 which controls the image duty cycle for the motion blur processing.

The common use of temporal undersampling virtually ensures some degree of aliasing on images which have periodic motion or object interaction near the frame rate or its harmonics. There are also some pathological cases where image misrepresentation due to temporal aliasing can make motion vector analysis impossible. When such motion vectors are required for de-interlacing or other standards conversion involving temporal transcoding, such conversions become artifact prone due to incorrect motion analysis. Temporal aliasing is also exaggerated when temporal transcoding is employed without the use of motion vectors.

It is very likely that temporal aliasing artifacts will appear at conversion rate beat frequencies. Thus, it is necessary to use beat frequencies which are at rates as high as possible.

8.2.4 Summary of Temporal Rate Analysis

In a digital image architecture which is interoperable and scalable across industries and applications, it is necessary to accommodate displays which exceed 70 Hz in refresh rate.

A choice of either 72 or 75 Hz for the display refresh rate could be beneficial in both advanced television systems and computer uses, particularly since NTSC may, in the future, become obsolete.

It is likely that flat panel displays such as active matrix liquid crystal color displays may begin to take hold in the market. These devices do not have characteristic flicker. Thus, it is possible to update the image at low rates, such

as 24 fps when presenting film images or 25 or 30 frames fps when presenting television images, without resulting flicker.

It is possible that future display devices may be developed which maintain the display on an effectively indefinite pixel-by-pixel basis. These devices will be frame rate independent and will permit even greater efficiency in the distribution of visual communications.

8.3 Bibliography

Boynton, Robert M., "Human Color Perception," Chapter 8, *Science of Vision*, K.N. Leibovic, Editor, Springer-Verlag, New York, 1990.

Leibovic, K. N., "Perceptual Aspects of Spatial and Temporal Relationships," Chapter 6, *Science of Vision*, K.N. Leibovic, Editor, Springer-Verlag, New York, 1990.

Maguire, William, Naomi Weisstein, and Victor Klymenko, "From Visual Structure to Perceptual Function," Chapter 9, *Science of Vision*, K.N. Leibovic, Editor, Springer-Verlag, New York, 1990.

Symes, Peter D., "The Enhanced Viewing Experience – What Does It Take?," Proceedings of the 26th Annual SMPTE Advanced Television and Electronic Imaging Conference, SMPTE, White Plains, NY, 1992.

APPENDIX F

STATEMENT CONCERNING USER REQUIREMENTS FOR SOURCE CODING AND MULTIPLEXING FOR BROADCAST APPLICATIONS

Documents
CCIR Study Groups
Period 1990-1994

Doc: 11/SRG-UR/Temp/2
Date: (17 July, 1992)
Original: English

Subject: Special Rapporteur Group 11/User Requirements

STATEMENT CONCERNING USER REQUIREMENTS FOR SOURCE CODING AND MULTIPLEXING FOR BROADCAST APPLICATIONS

Introduction

It is widely recognized today that broadcasting is in the middle of a revolution leading towards the "all-digital" broadcast system of the future.

In addition to the established trend to all digital operation in television production and distribution, the development of systems for delivering digital television to the home via terrestrial, cable and satellite digital networks is progressing rapidly on a world-wide basis.

CCIR Activities

In recognition of this dominant trend in broadcast development and of the importance and relevance of the work of ISO/IEC/MPEG, the recent meetings (May 1992) or CCIR Study Groups 10, 11 and CMTT appointed a Special Rapporteur Group to draw up Applications Requirements for digital source coding and multiplexing and to prepare a Report on them for transmission to the appropriate groups in MPEG (see doc. 11/135 attached).

Considerable progress towards the definition of the requirements for certain broadcasting applications has already been made in the CCIR. See docs. 11-AHG-DC/16, 11-AHG-DC/17, 11-AHG-DC/29.

Of particular importance and the highest priority in the work of Study Group 11, at the present time, is the definition of system(s) for the broadcasting of digital television in narrowband (6,7,8 MHz) terrestrial channels. This work is proceeding in Task Group 11/3. In addition, accelerated studies are taking place to define system(s), having a maximum of commonality with these terrestrial system(s), for the broadcasting of digital television in cable and telecommunications channels and in broadcasting satellite channels. This work is proceeding in Task Group CMTT/2 and JWP-10-11/5 respectively.

While it is as yet too early to be precise about the detailed requirements for these services, it is appropriate at this time to draw attention to the great interest in systems able to support hierarchical source and channel coding to provide HDTV, EDTV and CDTV (STV) transmission and receiver options for a broadcast signal with transmission bit rates in the range 20-30 Mbit/s.

System Example

As an example of the sort of system under discussion, (purely for illustration), a hierarchical (scalable) Digital Terrestrial Television Broadcast (DTTB) system might provide an HDTV quality at 24 Mbit/s, with a nested EDTV quality of 12 Mbit/s and a core CDTV quality at 4 Mbit/s. A common transport structure would be used for all delivery systems.

Timescale

In recognition of the rapid time-scale envisaged for standardisation within MPEG, the Special Rapporteur Group is charged with the production of its first Report concerning the requirements for broadcasting and secondary distribution in September 1992. Further Reports will be produced as studies progress, on this and other areas of interest, such as recording.

The schedule of meetings envisaged for CCIR Study Group 11, relevant to this work, is as follows:

-Task Group 11/4, Haromisation	Oct 13-15, 1992
-Task Group 11/1 HDTV Production	Dec 14-18, 1992
-Task Group 11/2 HDTV Interfaces	Dec 14-18, 1992
-Task Group 11/3 DTTB	Dec 14-18, 1992

The participation of an MPEG liaison representative in these meetings in invited.

It is envisaged that a more comprehensive second Report will be submitted to MPEG early in 1993.

APPENDIX G

TASK GROUP 11/4

PROPOSAL FOR A DRAFT NEW RECOMMENDATION

HARMONISATION OF DIGITAL METHODS FOR DELIVERY SYSTEMS FOR TELEVISION SERVICES TO THE HOME

Documents
CCIR Study Groups
Period 1990-1994

Doc:TG-11/4-Temp 05 Rev.1
Date: 15 Oct, 1992
Original: English

Subject: Question 119/11

Task Group 11/4

Proposal for a draft new Recommendation

Harmonisation of Digital Methods for Delivery Systems for Television Services to the Home

The CCIR

considering

(a) the rapid development of digital methods for the delivery of television and other image services to the home over terrestrial broadcast, cable, satellite channels and by way of pre-recorded media;

(b) the differing characteristics and capabilities of such delivery methods;

(c) the need for services to be available at differing levels of quality for diverse applications;

(d) that consumer receiving equipment is likely to exist at differing levels of capability;

(e) the potential advantages of a coordinated approach to the development of such delivery systems:

RECOMMENDS

that in the development of Recommendations and standards concerning digital methods of delivery for television and image services to the home, consideration be given to the following principles for their harmonisation:-

(a) that the source coding should be based on common processing algorithms and have a maximum of shared parameters;

(b) that hierarchical (scalable) coding, with a number of embedded levels of quality, be employed to allow the use of receivers having differing levels of capability;

(c) that delivery systems use a flexible, dynamically-variable multiplex of service elements;

(d) that headers/descriptors be included in the data stream to enable the receiver to identify and process a range of services that may have differing characteristics.

---

* The Director of the CCIR is requested to bring this Recommendation to the attention of the IEC, the ISO and the CCITT.